A study of developmentally regulated genes in
Physarum polycephalum

Thesis Submitted for the Degree
of Doctor of Philosophy at the
University of Leicester

By

Emma Swanston

Department of Genetics
University of Leicester

September 2000
A study of developmentally regulated genes in *Physarum polycephalum*

by Emma Swanston

Abstract

In the life cycle of the Protist, *Physarum polycephalum*, uninucleate amoebae develop into multinucleate, syncytial plasmodia via an irreversible developmental transition. Previously a subtracted cDNA library was screened and a novel class of genes was identified that were expressed primarily during development. These genes were called red genes (regulated in development).

My first aim was to identify and characterise more red genes. I isolated two partial cDNA clones, D11/1100 (redE) and A18/1020 (mynD). Northern blotting showed that both of the genes represented red genes; for redE this was confirmed by RT-PCR. Southern blotting showed that redE is a single-copy gene, while mynD is a single-copy gene that is a member of a gene family. A18/1020 was renamed mynD (myosin developmental) because database searches showed that the cDNA encodes part of the tail domain of a type II myosin heavy-chain protein. The complete redE coding sequence, plus some of the upstream promoter sequence was obtained by cloning a genomic restriction fragment. Database searches showed that the deduced RedE protein was not homologous to any known proteins, although it contained several potential phosphorylation and glycosylation sites. Thus, the function of RedE is unknown. As a first step towards investigating red gene function, I attempted to make a vector that would allow the manipulation of gene expression.

My second aim was to examine the replication of the red genes in the plasmodial S-phase, because genes that are expressed at low levels in the plasmodium had not previously been studied. Using density shift and gene dosage experiments, I showed that redE is replicated in early S-phase, and that two previously discovered red genes, redA and redB, are replicated in mid and early S-phase respectively. Two-dimensional gel electrophoresis showed that both redB and redE are flanked by replication origins that are activated at the onset of S-phase in the plasmodium.
ACKNOWLEDGEMENTS

I want to thank Juliet Bailey for the support and encouragement that she has provided over the past four years. I also thank Jenny Dee, Jenny Foxon, Rick Kilmer-Barber, Lynn Cook, Kathryn Evans and Angela Thomas for their endless help and acknowledge two undergraduate project students, Lorraine King and Aki Kobayashi for their contribution to the work in Chapter 4. I am also grateful to Laszlo Nyitray for sharing unpublished findings.

I want to thank Gérard Pierron for his supervision and kindness. I also thank Marianne Bénard and Yvette Florentin for sharing their lab expertise and acknowledge Chrystelle Maric for, er, something! I am grateful to the family of Gérard for making me so welcome in Villejuif.

Finally I want to thank my family and friends, especially Mum, Dad, Amy, Laura, Charley, Gareth, Rosie, Wendel, Kezzor and Nics.
TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .......................................................................................1
  1.1 THE P. POLYCEPHALUM LIFE CYCLE .............................................................1
  1.1.1 Amoebae ........................................................................................................1
  1.1.2 Plasmodia ........................................................................................................2
  1.1.3 Heterothallic development .............................................................................3
  1.1.4 Apogamic development ................................................................................4
  1.1.5 The amoebal-plasmodial transition in heterothallic and apogamic development ..................................................4
  1.2 THE CYTOSKELETON ....................................................................................6
     1.2.1 Microtubule organisation ..........................................................................6
     1.2.2 Microfilament organisation .......................................................................8
  1.3 STAGE SPECIFIC GENES AND PLASMODIAL DEVELOPMENT ......................10
     1.3.1 Differential gene expression in amoebae and plasmodia ..........................10
     1.3.2 Developmental mutants ..........................................................................12
     1.3.3 Constructing and screening the subtracted cDNA library ..........................13
  1.4 DNA TRANSFORMATION IN P. POLYCEPHALUM ..........................................19
  1.5 THE P. POLYCEPHALUM GENOME ..............................................................21
  1.6 DNA REPLICATION .......................................................................................21
     1.6.1 DNA replication in P. polycephalum .........................................................23
  1.7 AIDS .............................................................................................................26

CHAPTER 2: MATERIALS AND METHODS ..................................................................28
  2.1 P. POLYCEPHALUM CULTURE AND TRANSFORMATION .............................28
     2.1.1 Gene loci and strains ..............................................................................28
     2.1.2 Media ........................................................................................................28
     2.1.3 Excysting and encysting amoebae .............................................................29
     2.1.4 Axenic culture of amoebae ......................................................................29
     2.1.5 Amoebal transformations ........................................................................30
     2.1.6 Inducing plasmodium development ........................................................32
     2.1.7 Axenic culture of plasmodia ....................................................................32
  2.2 NUCLEIC ACID ISOLATION ...........................................................................33
     2.2.1 Quantification of nucleic acids by spectrophotometry ............................33
     2.2.2 RNA isolation ..........................................................................................33
     2.2.3 Isolation of genomic DNA from axenically grown amoebae ..................35
     2.2.4 Isolation of genomic DNA from microplasmodia ....................................36
     2.2.5 Isolation of plasmid DNA ......................................................................36
     2.2.6 Rapid isolation of plasmid DNA .............................................................37
  2.3 DNA CLONING ............................................................................................37
     2.3.1 Restriction enzyme digestion ..................................................................37
     2.3.2 Agarose gel electrophoresis ....................................................................38
     2.3.3 Isolating restriction fragments of interest ...............................................38
     2.3.4 Vector and insert preparation ..................................................................39
     2.3.5 DNA ligation and bacterial transformation .............................................39
     2.3.6 Making frozen stocks of bacterial cultures ..............................................40
  2.4 NUCLEIC ACID ANALYSIS ..........................................................................40
     2.4.1 Southern blotting .....................................................................................40
     2.4.2 Northern blotting .....................................................................................41
     2.4.3 Colony blots .............................................................................................41
     2.4.4 Radio-labelling probes and hybridisation .................................................42
     2.4.5 DNA sequencing .....................................................................................42
     WORK CONDUCTED IN VILLEJUIF .................................................................43
  2.5 P. POLYCEPHALUM CULTURE AND NUCLEIC ACID ISOLATION .................43
     2.5.1 Culturing macroplasmodia ......................................................................43
     2.5.2 Treating a plasmodium with Bromo-deoxyUridine (BrdU) .....................44
     2.5.3 Isolation of genomic DNA from macroplasmodia ..................................44
     2.5.4 RNA isolation and RT-PCR .................................................................45
  2.6 RESTRICTION ENZYME DIGESTION AND AGAROSE GEL ELECTROPHORESIS ......................................................................................46
     2.6.1 Restriction enzyme digestion ..................................................................46
     2.6.2 Agarose gel electrophoresis ....................................................................47
     2.6.3 Isolating restriction fragments of interest ...............................................47
  2.7 NUCLEIC ACID ANALYSIS ..........................................................................48
     2.7.1 Southern blotting .....................................................................................48
     2.7.2 Separating the BrdU-substituted non-BrdU substituted DNA ..................48
     2.7.3 2D gel electrophoresis .............................................................................49
     2.7.4 Radio-labelling probes and hybridisation .................................................49
CHAPTER 1: INTRODUCTION

*Physarum polycephalum* is an acellular slime mould that is classified as a Protist within the division Mycetozoa. The Mycetozoa are separated into three main groups: (i) the acellular Myxogastrid slime moulds, such as *P. polycephalum* (ii) the cellular Dictyostelid slime moulds, such as *Dictyostelium discoideum*, and (iii) the Protostelid slime moulds (reviewed in Baldauf and Doolittle, 1997). Molecular phylogenies derived from the analysis of actin, β-tubulin and elongation factor-1α genes have shown that the Mycetozoa represent a closely related division that are more related to fungi and animals than to plants (Baldauf and Doolittle, 1997). *P. polycephalum* is generally found in temperate vegetative areas.

1.1 THE *P. POLYCEPHALUM* LIFE CYCLE

*P. polycephalum* has two vegetative growth phases within its life cycle; individual cells can exist as haploid uninucleate amoebae or multinucleate syncytial plasmodia. These two very different cell types are linked by an irreversible developmental transition called the amoebal-plasmodial transition (APT).

1.1.1 Amoebae

*P. polycephalum* exists predominantly as haploid amoebae which are 10-20μm in diameter and live in the soil. They feed by phagocytosis on microbes and spores and move by amoeboid locomotion. In favourable conditions, amoebae divide by mitosis with cytokinesis to produce two identical daughter cells. The nuclear envelope dissociates during mitosis and asters can be visualised at the spindle poles in the ‘open’ mitosis that is characteristic of amoebae (Havercroft and Gull, 1983). When two genetically different amoebae meet they fuse and develop into a diploid plasmodium (Fig. 1.1).

In adverse conditions, amoebae can undergo one of two reversible developmental transitions. In very wet conditions they develop into flagellates which have one long and one short flagellum (Fig. 1.1). Flagellates lose some of the basic functions of amoebae; they are unable to feed and undergo mitosis. These cells use their flagella to swim to dryer regions where they absorb their flagella and revert back to amoebae (Glyn and Gull, 1990). If amoebae starve or become cold, they develop into cysts which have a resistant cell wall (Fig. 1.1). These cysts can lie dormant in the soil for many months. Once favourable conditions prevail, excystment occurs to release active amoebae.
1.1.2 Plasmodia

Plasmodia are giant cells that live on the surface of soil, leaf litter or rotting wood. They feed predominantly by phagocytosis on microbes and spores, although they also secrete digestive enzymes and take up the products by pinocytosis. Plasmodia have an intricate vein network through which protoplasmic streaming occurs, thereby providing a basis for movement. The many nuclei of a plasmodium divide by mitosis without cytokinesis, which leads to the formation of a syncytium; a typical plate grown plasmodium can contain as many as $10^8$ nuclei. The nuclear envelope remains intact during mitosis and asters are not visualised at the spindle poles in the ‘closed’ intranuclear mitosis that is characteristic of plasmodia (Havercroft and Gull, 1983). Plasmodia can grow to cover an area of several square metres by undergoing mitosis without cytokinesis, or by fusing with other plasmodia. Plasmodial fusion is controlled by the fus (fusion) genes. There are three fus loci, fusA, fusB and fusC and each has two known alleles (reviewed in Haugli et al., 1980). The two fusA alleles, fusA1 and fusA2 are co-dominant; fusion is only permitted between plasmodia that are identical at the fusA locus. While the fusB2 allele is dominant to fusB1 allele, and the fusC2 allele is dominant to fusC1 allele; fusion is permitted between plasmodia that each contain at least one dominant allele but not between a plasmodium containing only recessive alleles and a plasmodium containing at least one dominant allele. If plasmodia are compatible at the fus loci, they fuse to produce a heterokaryon (Haugli et al., 1980).

Studies involving the incorporation of radioactive precursors into replicating DNA have shown that the mitotic cycles in the many nuclei of the plasmodium are synchronous. When a plasmodium has enough food and is incubated at optimal temperatures, the nuclei divide synchronously approximately every 10 hours (Nygaard et al., 1960; Braun et al., 1965). Flow cytometry of Hoeschst-stained nuclei has since shown that 99% of the nuclei enter and exit mitosis simultaneously (Kubbies and Pierron, 1983). During the incorporation studies, Nygaard et al. (1960) noticed that DNA replication begins directly after mitosis and nuclear division; there is no G1-phase in the plasmodium (Braun et al., 1965). Thus, the start of S-phase is defined as the observation of telophase. The 3 hour S-phase is followed by a 6-7 hour G2-phase which leads directly to the next mitosis (Braun et al., 1965).

When a plasmodium starves in the dark, it develops into a sclerotium with a thick cell wall (Fig. 1.1). A sclerotium has groups of spherules which contain up to 14 nuclei and can lie dormant in the soil for many years (Gorman and Wilkins, 1980). Once favourable conditions prevail, the sclerotium reverts back to an active plasmodium. When a plasmodium starves in the light, sporangia form and meiosis occurs in the spores (Fig. 1.1). Of the four meiotic products, three degenerate and one survives to develop a thick
cell wall (Laane and Haugli, 1976). In favourable conditions, the spores germinate to produce haploid amoebae, and the life cycle begins again (Fig. 1.1).

Daniel and Rusch, (1961) first cultured plasmodia in shaken axenic media. In shaken culture a macroplasmodium becomes fragmented. Although the mitotic cycles in the nuclei of a single small ‘microplasmodium’ are synchronous, there is no synchrony between the different microplasmodia in the culture. When a small volume of axenic culture is transferred to suitable media, the microplasmodia fuse to form a macroplasmodium. After fusion, the cell cycles of the many nuclei quickly synchronise.

1.1.3 Heterothallic development
Natural isolates of *P. polycephalum* undergo heterothallic development; two amoebae fuse and develop into a diploid plasmodium. This process is controlled by three unlinked mating type loci, *matA*, *matB* and *matC*. The frequency of amoebal fusion events is influenced by *matB* and *matC*. Amoebae that are heteroallelic for *matB* are 100-1000 times more likely to fuse than amoebae that are homoallelic for *matB* (Youngman *et al.*, 1981). Amoebae that are heteroallelic for *matC* can fuse at higher pH than amoebae that are homoallelic for *matC* (Kawano *et al.*, 1987). The *matA* locus has no effect on amoebal fusion, but controls the development of the fusion cell. If amoebae that are heteroallelic for *matA* fuse, the haploid nuclei fuse to produce a diploid zygote which subsequently develops into a diploid plasmodium (Youngman *et al.*, 1981). In contrast, if amoebae that are homoallelic for *matA* fuse, plasmodium development is blocked (Youngman *et al.*, 1981) and the fusion cell usually splits apart to produce two uninucleate, haploid amoebae (Bailey *et al.*, 1990). Thus far, 18 *matA* alleles, 15 *matB* alleles and 3 *matC* alleles have been identified (reviewed in Barber, 1998).

1.1.4 Apogamic development
Spontaneous mutations in laboratory strains of *P. polycephalum* produced new strains that are able to undergo apogamic development; a single haploid amoeba develops into a haploid plasmodium. The resulting haploid plasmodia are capable of sporulation, although the spore viability is very low. It is thought that the few viable spores arise from small populations of diploid nuclei that are found within the plasmodium (Laffler and Dove, 1977). Apogamic strains of *P. polycephalum* carry a mutation at the *gad* (greater asexual differentiation) locus. Thus far, 29 *gad* mutations have been identified. Adler and Holt (1977) and Gorman *et al.* (1979) showed that 28 of the *gad* mutations are genetically inseparable from *matA*. This *matA*-inseparable class of mutations have been designated *gadA* (Anderson and Hutchins, 1986). The apogamic development of *gadA* mutants is generally inhibited at high temperatures; CL and LU352 amoebae, which are both *matA2*
Apogamic amoebae can lose the ability to develop into plasmodia or 'self' following mutagenesis. The isolation of non-selfing strains led to the identification of many mutations that block apogamic development, but have no affect on amoebal growth (Wheals, 1973; Anderson and Dee, 1977; Solnica-Krezel et al., 1995). Most of these npf (no plasmodium formation) mutations map to the matA locus (Anderson et al., 1989). The analysis of strains carrying matA-linked mutations, showed that the mutations are located in one of two complementation groups, called npfB and npfC (Anderson et al., 1989). The physical association of the gadA, npfB and npfC mutations with matA, shows that matA is a complex multifunctional locus (Anderson et al., 1989; Bailey, 1995). The cloning of matA will facilitate a further understanding of its function. The analysis of strains which carry npf mutations which do not map to matA are discussed further in section 1.3.2.

Apogamic strains of *P. polycephalum* are valuable to scientists because they remain haploid throughout their life cycle. This allows the phenotype of recessive mutations that affect plasmodium development to be identified and studied easily. These strains are also useful for gene disruption studies because if known genes are knocked out, the phenotype of the mutant strain can be studied throughout its life cycle.

### 1.1.5 The amoebal-plasmodial transition in heterothallic and apogamic development

Proliferating apogamic and heterothallic amoebae only form plasmodia when their density reaches a critical level. Youngman *et al.* (1977) showed that dense cultures of apogamic amoebae induced plasmodium development in sparse amoebal cultures from which they were separated by a filter. Shipley and Holt (1982) and Nader *et al.* (1984) further showed that dense cultures of heterothallic amoebae induced fusion competence, and therefore plasmodium development, in sparse amoebal cultures. These observations led the authors to conclude that a diffusible inducer accelerates the mating of heterothallic amoebae and the development of apogamic strains. Although the diffusible inducer has not been isolated, Nader *et al.* (1984) isolated an inducer of plasmodium development from another Myxogastrid slime mould, *Didymium iridis*, and showed that the 120kDa glycoprotein could also induce plasmodium development in *P. polycephalum*.

Time-lapse cinematography and immunofluorescence microscopy have been used to study cellular events in the amoebal-plasmodial transition in both heterothallic and apogamic plasmodium development (Bailey *et al.*, 1987 and 1990; Solnica-Krezel *et al.*, 1990 and 1991). During heterothallic development, amoebae fuse in pairs and form binucleate fusion cells. Approximately two hours later, the nuclei fuse to form a diploid zygote which enters
an extended cell cycle which is approximately 2.3 times longer than a normal amoebal cell cycle (Bailey et al., 1990). At the end of the extended cell cycle the cell undergoes ‘closed’ mitosis without cytokinesis to form a binucleate cell. The binucleate cell then enters a shortened cell cycle that is approximately 0.7 times as long as a normal amoebal cell cycle (Bailey et al., 1990). The cell growth proceeds at a constant rate throughout the extended cell cycle, which results in a binucleate cell with a high cytoplasm:DNA ratio. The subsequent shortened cell cycle returns this ratio to normal levels (Bailey et al., 1990). Subsequent rounds of mitosis without cytokinesis and plasmodial fusion events form a macroplasmodium that contains millions of diploid nuclei (Bailey et al., 1990). The cell becomes committed to plasmodium development and loses the ability to flagellate at around the time of nuclear fusion (Shipley and Holt, 1982; Bailey et al., 1990). The cell also begins to acquire plasmodial characteristics during the extended cell cycle, such as the ability to ingest amoebae and fuse with other plasmodia (Bailey et al., 1990).

As in heterothallic development, during apogamic development an amoeba enters an extended cell cycle that is approximately 2.3 times longer than a normal amoebal cell cycle (Fig. 1.2). After the subsequent 'closed' mitosis without cytokinesis the cell enters a cell cycle that is 0.7 times the length of a normal amoebal cell cycle (Fig. 1.2; Bailey et al., 1987). Subsequent rounds of mitosis without cytokinesis and plasmodial fusion events form a macroplasmodium containing millions of haploid nuclei. The cell becomes committed to plasmodium development and loses the ability to flagellate approximately halfway through the extended cell cycle (Fig. 1.2; Blindt et al., 1986; Bailey et al., 1987). As in heterothallic development, the cell begins to acquire plasmodial characteristics during the extended cell cycle (Bailey et al., 1987).

The cellular events of the amoebal-plasmodial transition are very similar in both heterothallic and apogamic development. The only real difference is in the timing of commitment and the loss of the ability to flagellate. In heterothallic development, commitment occurs up to 20 hours before the first plasmodial mitosis and the cells are still capable of flagellation at, or just after, becoming committed to development (Bailey et al., 1990). In contrast, in apogamic development commitment occurs up to 13.5 hours before the first plasmodial mitosis (Bailey et al., 1987; Bailey et al., 1990) and the cells lose the ability to flagellate 3-5 hours before commitment (Blindt et al., 1986; Bailey et al., 1987). Commitment was defined using replating assays (section 2.2.2). In this procedure, amoebae were plated onto several plates and the plates were incubated at 26°C. At various time points, the cells on a single plate were harvested and plated onto several fresh plates at low density to remove the cells from the presence of the diffusible inducer (see above). After 2-3 days the committed cells formed small plasmodia, and after 4-5 days the non-committed cells formed amoebal colonies (Bailey et al., 1987). As commitment is
operationally defined, it is not clear whether an apogamic committed cell is at the same developmental stage as a committed heterothallic cell.

Cells become committed to plasmodium development, lose the ability to flagellate and gradually acquire plasmodial characters during the extended cell cycle. It is a key time in development since the changes in gene expression and cellular organisation that result in plasmodium formation are initiated at this time. This is further discussed in sections 12.1, 1.1.2.2 and 1.3.1.

1.2 THE CYTOSKELETON

1.2.1 Microtubule organisation

Microtubules are structural filaments that constitute a major part of the cytoskeleton in eukaryotic cells. They are key to a variety of cellular functions including intracellular transport and the formation of the mitotic spindle and flagella. Protofilaments are built of tubulin subunits which are heterodimers of $\alpha$ and $\beta$ tubulin, and 13 protofilaments wrap around a hollow core to form a microtubule (Alberts et al., 1989). There are at least four $\alpha$ tubulin loci ($altA$, $altB$, $altC$ and $altD$) and three $\beta$ tubulin loci ($betA$, $betB$ and $betC$) in $P.\ polycephalum$ (Schedl et al., 1984). Studies have shown that the $altB$ locus comprises two tightly linked $\alpha$ tubulin genes called $altB(N)$ and $altB(E)$ (Monteiro and Cox, 1986; Schedl et al., 1984). The tubulin genes encode several distinct tubulin polypeptides; altogether four $\alpha$ tubulin isotypes ($\alpha1A$, $\alpha1B$, $\alpha2B$ and $\alpha3$) and three $\beta$ tubulin isotypes ($\beta1A$, $\beta1B$ and $\beta2$) have been identified (reviewed in Burland et al., 1993b). The $\alpha1A$, $\alpha3$, $\beta1A$ and $\beta1B$ tubulin isotypes are found in amoebal microtubules, while the $\alpha1A$, $\alpha1B$, $\alpha2B$, $\beta1B$ and $\beta2$ tubulin isotypes are found in plasmodial microtubules (Burland et al., 1993b). Thus, the $\alpha3$ and $\beta1A$ tubulin isotypes are amoeba-specific and the $\alpha1B$, $\alpha2B$ and $\beta2$ tubulin isotypes are plasmodium-specific.

Immunofluorescence microscopy with isotype specific anti-tubulin antibodies has been used to study the arrangement of microtubules in amoebae and plasmodia (Havercroft and Gull, 1983; Salles-Passador et al., 1991). Interphase amoebae contain a single microtubule organising centre (MTOC) from which a dense network of microtubules radiates (Havercroft and Gull, 1983; Bailey et al., 1990). The MTOC is closely associated with a pair of centrioles and linked to the nucleus. During early prophase the MTOC duplicates and divides. The two daughter MTOCs then migrate to opposite sides of the cell and nucleate the compact spindle that is visualised at metaphase (Havercroft and Gull, 1983). The breakdown of the nuclear envelope at metaphase permits the formation of asters that
radiate from the MTOC to the cell surface (Havercroft and Gull, 1983). The nuclear envelope reforms after mitosis and cytokinesis produces two identical daughter cells. The centrioles act as the basal bodies of the two flagella that form when amoebae develop into flagellates, which explains why flagellates cannot undergo mitosis (Havercroft and Gull, 1983).

Plasmodia contain a dense network of cytoplasmic microtubules that do not have an organised orientation, except close to the veins where they are located parallel to the long axis of the vein (Salles-Passador et al., 1991). These cytoplasmic microtubules are not nucleated by obvious MTOCs (Salles-Passador et al., 1991) and the nuclei of interphase plasmodia do not contain an MTOC (Havercroft and Gull, 1983). A single plasmodial MTOC appears in the nucleus during early prophase and divides. The two daughter MTOCs then migrate to opposite sides of the nucleus and nucleate the mitotic spindle (Havercroft and Gull, 1983). Asters are not formed during mitosis because the nuclear envelope stays intact (Havercroft and Gull, 1983). Cytokinesis is absent leading to the formation of a syncytium (Havercroft and Gull, 1983).

Immunolocalisation experiments showed that the amoeba-specific $\alpha 3$ tubulin is gradually lost from cells during the APT, while the plasmodium-specific $\beta 2$ tubulin is gradually acquired by cells (Diggins-Gilicinski et al., 1989; Solnica-Krezel et al., 1990 and 1991). The disappearance of the $\alpha 3$ tubulin isotype starts in the extended cell cycle of uninucleate developing cells and is finished by the quadrinucleate stage. In contrast the acquisition of $\beta 2$ tubulin starts during the extended cell cycle of uninucleate developing cells and continues through development (Solnica-Krezel et al., 1990). Thus, the extended cell cycle is when the changes in the organisation of the cytoskeleton are initiated. An examination of $\beta 2$ tubulin positive cells showed that cells do not always enter the closed mitosis that is characteristic of plasmodia at the end of the extended cell cycle (Solnica-Krezel et al., 1991). Some cells entered the open mitosis that was characteristic of amoebae, while others entered a chimeric mitosis that had features of both the open mitosis that is characteristic of amoebae and the closed mitosis that is characteristic of plasmodia. Cells undergoing chimeric mitosis contained a plasmodial MTOC that was $\beta 2$ tubulin positive, in addition to an amoebal cytoplasmic MTOC that was $\alpha 3$ tubulin positive, which suggests the plasmodial MTOC is not derived from the amoebal MTOC (Solnica-Krezel et al., 1991). The studies identified $\alpha 3$ tubulin positive cells that were undergoing the closed mitosis that is characteristic of plasmodia and $\beta 2$ tubulin positive cells that were undergoing the open mitosis characteristic of amoebae (Diggins-Gilicinski et al., 1989;
Solnica-Krezel et al., 1990 and 1991). Thus, the presence of α3 tubulin or β2 tubulin alone does not determine whether the cell undergoes an amoebal or plasmodial mitosis respectively.

1.2.2 Microfilament organisation

Actin filaments or microfilaments are structural filaments that form a major part of the cytoskeleton in eukaryotic cells. They are key to a variety of cellular functions including cytokinesis and cell motility. Microfilaments are formed by the polymerisation of globular-actin (G-actin) into filamentous-actin (F-actin; Alberts et al., 1989). There are four unlinked actin loci (ardA, ardB, ardC and ardD) in *P. polycephalum* (Schedl and Dove, 1982). Mendelian analysis has shown that ardB, ardC and ardD loci are comprised of a single gene, while the ardA locus is more complex and is comprised of two tightly linked actin genes which were designated ardA and ardE (Schedl and Dove, 1982; Hamelin et al., 1988; Pallotta et al., 1989). The cloning and sequencing of ardA, ardB and ardC cDNAs (Hamelin et al., 1988) showed that the genes encode the same protein; the ardB and ardC genes are 99% identical, while the ardA gene is 93% identical to both the ardB and ardC genes (Hamelin et al., 1988). Northern blotting analysis has shown that the ardB and ardC transcripts are present in all stages of the *P. polycephalum* life cycle (Hamelin et al., 1988). In contrast, the ardA mRNA could not be detected at any stage, although very low levels of transcript have since been detected in microplasmodia (Arellano et al., 1992). The ardB and ardC mRNAs are not equally abundant; S1 nuclease protection and dot hybridisation assays have shown that there is a 3:1 ratio of ardC:ardB transcript in amoebae and plasmodia (Hamelin et al., 1988). Three major bands corresponding to ardA, ardB and ardC are visualised when a Southern blot is probed with the ardC cDNA at high stringency, which shows that ardD and ardE are less homologous to the other actin genes (Hamelin et al., 1988). The cloning and sequencing of the ardD cDNA showed that the deduced ArdD protein is 84% identical to the ArdA, ArdB and ArdC proteins (Adam et al., 1991). Northern blotting analysis has shown that ardD is not expressed in amoebae, but is expressed in plasmodia and spherules (Adam et al., 1991).

In amoebae, most of the actin is found in a continuous cortical layer of microfilaments that are associated with the plasma membrane (Clayton et al., 1983), although microfilaments are also concentrated within broad pseudopodia (reviewed in Stockem and Brix, 1994). Ultrastructural studies have shown that the amoebal microfilament system is a three-dimensional network of microfilaments (Pagh et al., 1985). In plasmodia, actin is arranged into either (i) a continuous cortical layer of microfilaments that is closely associated with the plasma membrane, or (ii) a dense fibrillar microfilament system (Stockem and Brix, 1994). The fibrillar microfilament network is unordered near the cell periphery but
becomes more ordered towards the centre of the cell, where groups of fibrils form helically twisted bundles that encircle the veins (Stockem and Brix, 1994).

It is generally accepted that the cellular arrangement of microfilaments is controlled by actin binding proteins (ABPs) that perform a variety of functions, such as regulating the formation of F-actin or cross-linking microfilaments (reviewed in Stockem and Brix, 1994; Hatano, 1994). Two P. polycephalum ABPs have been analysed in detail; profilin is a monomer sequestering protein which regulates the assembly of F-actin by sequestering G-actin (Schlüter et al., 1997), while fragmin is an actin severing protein that can sever F-actin and cap its ends to prevent microfilament growth (Hatano, 1994). Studies have shown that amoebae and plasmodia posses cell-type specific profilin and fragmin isotypes. The amoeba-specific profilin and fragmin isotypes, ProA and FrgA, are found in amoebae, but not plasmodia, while the plasmodium-specific profilin and fragmin isotypes, ProP and FrgP, are found in plasmodia, but not in amoebae (Binette et al., 1990; T'Jampens et al., 1997 and 1999; Uyeda et al., 1988). ProA and ProP are 66% identical (Binette et al., 1990), while FrgA and FrgP are 65% identical (T'Jampens et al., 1997). The existence of cell-type specific profilin and fragmin proteins suggests that the isotypes perform different functions. Marcoux et al. (1999) recently provided evidence for this notion by overexpressing proA and proP in profilin deficient yeast cells. As the profilins did not correct the mutant phenotypes equally well, the authors conclude that proA and proP are not functionally equivalent (Marcoux et al., 1999).

Myosin is the most important actin binding protein (reviewed in Hatano, 1994; Stockem and Brix, 1994). Myosin molecules are subdivided into conventional myosin molecules and unconventional myosin molecules. Conventional myosin, or myosin II is found in skeletal muscle, smooth muscle and most non-muscle cells. These molecules are hexamers comprising two heavy chain proteins that are folded into a head and a tail domain and four light chain proteins that are non-covalently bound to the head domain of the heavy chain proteins (Hatano, 1994; Stockem and Brix, 1994). The myosin head domain has a strong and a weak actin binding site and acts as an Mg\(^{2+}\)-ATPase. The myosin head binds to F-actin and the hydrolysis of ATP provides energy which translocates the molecule along the microfilament. Generally, myosin II molecules provide the force-generating mechanism for microfilament contractile activities (Hatano, 1994; Stockem and Brix, 1994). Unconventional myosin, or myosin I is found in all eukaryotic cells. These molecules are dimers that have one heavy chain protein that is folded into a head and a short tail domain and one light chain protein that is non-covalently bound to the head domain (Hatano, 1994; Stockem and Brix, 1994). The variety in the tail domains of unconventional myosins allows the molecules to perform a variety of functions, such as cross-linking.
microfilaments, in addition to providing the force for microfilament contractile activities (Hatano, 1994; Stockem and Brix, 1994).

Myosin co-localises to the microfilament systems in amoebae and plasmodia and provides the force for contractile activities (Stockem and Brix, 1994). Ohl and Stockem (1995) examined the distribution of type II myosin in plasmodia and showed that the fibrils that encircle the veins have a higher type II myosin content than the cortical layer of microfilaments, and that type II myosin is essential for the vein contraction that results in protoplasmic streaming. Tanigushi et al. (1980) and Kohama et al. (1986) isolated type II myosin from amoebae and plasmodia and compared the two proteins. Electron microscopy showed that the purified myosin molecules looked the same and further analyses showed that the amoebal and plasmodial myosins were comprised of two heavy-chain proteins, two phosphorylatable light-chain proteins and two Ca^{2+}-binding light-chain proteins. Two-dimensional gel electrophoresis, peptide mapping and immunological studies showed that the heavy chain proteins and the phosphorylatable light chain proteins of amoebal and plasmodial myosin were different. In contrast, the Ca^{2+}-binding light chain proteins were common to amoebal and plasmodial myosin (Kohama et al., 1986).

Using immunofluorescence microscopy with isotype specific anti-tubulin antibodies, Uyeda et al. (1987) showed that the amoebal myosin is present in amoebae and uninucleate developing cells, but absent in plasmodia, while the plasmodial myosin is absent in amoebae, but present in uninucleate developing cells and plasmodia. The authors concluded that the ‘switch’ from the synthesis of amoebal myosin to the synthesis of plasmodial myosin occurs during the extended cell cycle of uninucleate developing cells (Uyeda et al., 1987). Further studies confirmed that amoebae and plasmodia have the same Ca^{2+}-binding light chain proteins; Kobayashi et al. (1988) isolated a cDNA representing the Ca^{2+}-binding light-chain protein and showed, by Southern blotting, that it represents a single-copy gene.

1.3 STAGE SPECIFIC GENES AND PLASMODIAL DEVELOPMENT
1.3.1 Differential gene expression in amoebae and plasmodia
The abundant proteins of amoebae and plasmodia were compared using two-dimensional gel electrophoresis (Turnock et al., 1981). The studies showed that while 76% of the proteins were common to both amoebae and plasmodia, the remaining 26% were cell-type specific; 12% were found only in amoebae, while 14% were found only in plasmodia. The authors suggested that a quarter of the differences were due to post-translational modifications. Thus, approximately 20% of the abundant proteins in *P. polycephalum* are cell-type specific (Turnock et al., 1981)
A further study by Pallotta et al. (1986) showed that amoebae and plasmodia have different populations of mRNA. The authors constructed cDNA libraries from amoebal and plasmodial RNA and screened each library with radio-labelled RNA from amoebae and plasmodia. Approximately 2.3% of the clones in the amoebal library hybridised only to the amoebal RNA, while 16% of the clones in the plasmodial library hybridised only to the plasmodial RNA (Pallotta et al., 1986). The most abundant cell-type specific clones were then analysed in more detail. The amoebal library contained only one abundant amoeba-specific clone which represented 1.2% of the library. A subsequent study showed that this cDNA encoded the amoebal-specific profilin isotype, ProA (Binette et al., 1990). As actin clones represented 4.2% of the library, actin was approximately 3.5 times more abundant than the most abundant amoeba-specific cDNA, proA (Pallotta et al., 1986). The plasmodial library contained four abundant clones which represented 4.8%, 2.4%, 1.4% and 1% of the library. Subsequent studies showed that the first three clones encoded proteins of unknown function (HapP: Lépine et al., 1995; Plasmin B: Laroche et al., 1989 and Plasmin C: Girard et al., 1990), while the fourth encoded the amoebal-specific profilin isotype, ProP (Binette et al., 1990). As actin clones represented 4.9% of the library, actin was slightly more abundant than the most abundant plasmodium-specific cDNA, hapP (Pallotta et al., 1986).

A second study by Sweeney et al. (1987) confirmed and extended the work of Pallotta et al. (1986). Again, the authors constructed cDNA libraries from amoebal and plasmodial RNA and screened each library with radio-labelled cDNA that was generated from amoebal and plasmodial RNA. Approximately 5% of the clones in the amoebal cDNA library hybridised only to amoebal RNA, while 5% of the clones in the plasmodial library hybridised only to plasmodial RNA (Sweeney et al., 1987). A sample of the amoeba-specific and plasmodium-specific clones were used to probe northern blots containing RNA from amoebae, plasmodia and cultures that were enriched for uninucleate developing cells, binucleate cells and quadrinucleate cells. The mRNAs of the amoeba-specific clones were detected in the amoebal RNA and to a lesser extent in the RNA representing the various developing stages, but not in plasmodia (Sweeney et al., 1987). The mRNA of the plasmodium-specific clones was first detected in the RNA representing uninucleate developing cells. In some cases the mRNA levels rose throughout development and reached high levels in the plasmodial RNA. In others, the mRNA levels remained very low throughout development and only reached high levels in the plasmodial RNA (Sweeney et al., 1987). These studies showed that the loss of the amoeba-specific transcripts and the gain of the plasmodium-specific occurs over several nuclear divisions, and that these changes are initiated during the extended cell cycle of the APT. These findings are in agreement with the work of Solnica-Krezel et al. (1988 and 1990) which showed the
disappearance of the amoeba-specific α3 tubulin, and the appearance of the plasmodium-specific β2 tubulin is initiated during the extended cell cycle of the APT and occurs over several nuclear divisions.

1.3.2 Developmental mutants
As discussed in section 1.1.4, the isolation of non-selfing apogamic amoebae, following mutagenesis, led to the identification of many mutations that block apogamic development, but do not generally affect amoebal growth (Wheals, 1973; Anderson and Dee, 1977; Solnica-Krezel et al., 1995). Studies have shown that although most of these npf (no plasmodium formation) mutations are linked to matA (section 1.1.4), some are not. The npf mutations that are not linked to matA are discussed in this section. Bailey et al. (1992a and b) and Solnica-Krezel et al. (1995) analysed six of the matA unlinked npf mutations in detail. In these studies, several markers were used to assess the time at which development was arrested: (i) the presence of the α3 tubulin and the β2 tubulin isotypes (ii) the ability to flagellate (iii) the number of nuclei the cells contained and (iv) the type of mitosis the cells underwent. The authors found that the matA-unlinked mutations could be subdivided into two categories: (i) those that block development during the extended cell cycle and (ii) those that block development after the extended cell cycle (Bailey et al., 1992a and b; Solnica-Krezel et al., 1995). These are discussed in turn.

The early acting npf mutations (npfA1 and npfG1) block apogamic development during the extended cell cycle. Strains carrying these mutations showed little sign of development and binucleate cells were never observed. The cells retained the ability to flagellate and remained α3 tubulin positive and β2 tubulin negative (Bailey et al., 1992a; Solnica-Krezel et al., 1995).

The late acting npf mutations (npfF1, npfK1, npfM1 and npfL1) block apogamic development after the extended cell cycle, and the exact time at which development is arrested depends on the mutation. Strains carrying these mutations showed distinct signs of development and binucleate cells were always observed (Bailey et al., 1992a and b; Solnica-Krezel et al., 1995). The cells passed through the extended cell cycle and, in most cases, became binucleate after undergoing plasmodial mitosis. Most of the cells lost the ability to flagellate and became α3 tubulin negative and β2 tubulin positive during development. In all cases, the cells began to appear and act abnormally at the extended cell cycle. The specific phenotype of the cells depended on the mutation that was carried (Bailey et al., 1992a and b; Solnica-Krezel et al., 1995). These observations underline the importance of the extended cell cycle in plasmodium development.
In the mutant strains, some aspects of plasmodial development occurred normally, e.g. cells carrying the *npfK1* and *npfM1* expressed β2 tubulin and underwent plasmodial mitosis, but did not ingest amoebae or fuse with other cells. These observations led to the idea that plasmodium development is controlled by several independent regulatory pathways that are activated by *matA* (Bailey *et al.*, 1992a; Burland *et al.*, 1993b). The late acting *npf* mutations are probably located in essential plasmodium-specific genes since they first affect plasmodium development during the extended cell cycle and the phenotype becomes more abnormal throughout development. While the early acting *npf* mutations could be located in a specific set of regulatory genes that control the APT. In order to investigate the interaction between individual *npf* genes, Solnica-Krezel *et al.* (1995) constructed and analysed two double mutant strains. The authors found that the mutant phenotype of cells of the first strain was characteristic of both parent strains, while the mutant phenotype of cells of the second mutant strain was characteristic of only one of the parent strains. These observations confirmed the idea that plasmodium development is controlled by several regulatory pathways. Solnica-Krezel *et al.* (1995) conclude that the gene products of the first double mutant strain act in different developmental pathways, while the gene products of the second mutant strain act in the same developmental pathway. It is thought that *matA* initiates a cascade of gene action that results in the formation of a plasmodium (Bailey, 1995 and 1997).

In order to identify development specific genes, Bailey *et al.* (1992a) constructed a cDNA library from a CL developing culture which contained a high proportion of developing cells and enriched the library by subtraction for genes that were expressed during development. The subtraction step was included because it increases the chance of finding development specific genes that may be expressed at relatively low levels and only in some of the cells within a mixed population. The following section gives a brief account of how the subtracted cDNA library was made and describes the initial screening of the library. A detailed account of the library construction is published in Bailey *et al.* (1999).

### 1.3.3 Constructing and screening the subtracted cDNA library

#### 1.3.3.1 Constructing the cDNA library

The cDNA library was made from a developing culture of the apogamic strain CL, which contained a high proportion of developing cells (Table 2.1; section 2.2.2; Bailey *et al.*, 1999). Total RNA was isolated from the culture and then poly(A)+ RNA was isolated from the total RNA using an oligo(dT) column. The poly(A)+ RNA was then reverse transcribed to make cDNA using oligo(dT)NotI primers (Bailey *et al.*, 1992a). After second strand synthesis, EcoRI linkers were ligated onto the ends of the cDNAs. The cDNAs were then digested with EcoRI and NotI and directionally cloned into EcoRI NotI.
digested pBluescript II KS (Stratagene). The cloned cDNAs had the poly(A)+ tail adjacent to the NotI site, and the 5' end adjacent to the EcoRI site. The plasmids were then transformed into E. coli XL1-Blue to form Madison Library 8 (ML8). ML8 contained approximately $1 \times 10^7$ plasmid-carrying bacteria, with 95% of the plasmids containing cDNA inserts. The inserts ranged from 500-2000bp, and averaged approximately 900bp (Bailey et al., 1992a).

Approximately 30% of ML8 was then amplified to form Madison Library 8 Amplified (ML8A), within which 70% of the plasmids contained cDNA inserts (Bailey et al., 1999). The cDNAs representing genes that are abundantly transcribed in amoebae and/or plasmodia were then subtracted from ML8A to enrich the library for cDNAs representing development-specific (or) low abundance mRNAs. First, ML8A was transfected with an M13 helper phage that had a disabled replication origin. The proteins encoded by the helper phage recognised the pBluescript M13 replication origin and generated single stranded phagemids containing sequence complementary to mRNA. Then the subtraction was carried out by hybridising the single stranded phagemids to poly(A)+ RNA from amoebae and plasmodia. The non-hybridised phagemids were then separated from the hybridised phagemids and transformed into E. coli DH5αMCR to form Madison Library 8 Subtracted (ML8S) within which 15% of the plasmids contained cDNA inserts (Bailey et al., 1999). This is much less than in ML8A because the subtraction procedure did not remove plasmids without inserts.

1.3.3.2 The preliminary screening of ML8S
In order to isolate genes that were expressed primarily during development, ML8S was screened with three subtracted cDNA probes. The probes were, (A) an amoebal probe, (P) a microplasmodial probe and (D) a developing cell probe. In each case the first strand of the cDNA probes was synthesised from the relevant poly(A)+ RNA using reverse transcriptase. After synthesis the RNA was denatured and the subtraction procedure was carried out by hybridising the single-stranded cDNA probes to poly(A)+ RNA. The amoebal cDNA probe was subtracted against poly(A)+ RNA from amoebae, the plasmodial cDNA probe was subtracted against poly(A)+ RNA from plasmodia and the developing cell cDNA probe was subtracted against poly(A)+ RNA from amoebae and plasmodia (Bailey et al., 1999). The hybridisations were not allowed to go to completion so that the low abundance cDNAs would remain single stranded, while the high abundance cDNAs would become double stranded. The non-hybridised cDNAs were then separated from the hybridised cDNAs and the second strand of the single stranded cDNAs was synthesised. The subtraction step enriched the amoebal and plasmodial cDNA probes for low abundance transcripts by removing the more abundant cell-type specific transcripts,
and enriched the developing cell cDNA probe for development specific (or) low abundance transcripts by removing the more abundant transcripts that represented constitutively expressed, amoebal-specific or plasmodium-specific genes. The subtraction procedure improved the sensitivity of the screening since abundant transcripts, that are absent from ML8S, would have represented the large majority of the labelled probe.

In the first screen of ML8S, approximately $3 \times 10^4$ colonies were plated onto LB-amp plates and incubated at $37^\circ C$ overnight before colony lifts were carried out. Three identical filters were lifted from each original plate and one filter of each set was hybridised with each of the subtracted probes, A, D or P (Bailey et al., 1999). The results of the screening of each set of three filters with the subtracted probes were then compared. Colonies that hybridised to all three probes were eliminated because their inserts represented constitutively expressed genes. Colonies that hybridised to one or two of the probes were chosen for further analysis, and frozen stocks were made of each; 150 such colonies were identified (Bailey et al., 1999). Colonies that hybridised only to the A probe, or to both the A and the D probes were named A1, A2 etc. Colonies that hybridised only to the P probe, or to both the P and the D probes were named P1, P2 etc. Colonies that hybridised only to the D probe were named D1, D2 etc.

The replica plates contained densely plated bacteria, which meant that many of the 150 frozen stocks were comprised of a mixed population of cells; a second round screen was therefore necessary to identify individual colonies of interest. Thirty of the frozen stocks were plated at a density of approximately 20 colonies/plate and incubated (Bailey et al., 1999). As before, a set of three replica filters were made from each original plate, and one filter of each set was screened against each of the three subtracted probes A, D or P. Twelve individual colonies hybridised to just one or two of the probes; these were chosen for further analysis and plasmid DNA was extracted from each (Bailey et al., 1999). The plasmid DNAs were then digested to release the insert from the vector, and the inserts were gel purified.

The expression pattern of the 12 cDNA clones identified above was examined by northern blotting to see if any of them represented differentially expressed genes. The purified inserts were hybridised to northern blots containing $10 \mu g$ of total RNA isolated from CL amoebae, microplasmodia and a developing cell culture which contained a high proportion of developing cells (56%: Table 2.1; section 2.2.2; Bailey et al., 1999). The blots were probed at high stringency ($65^\circ C$). When the blots were analysed, three different hybridisation patterns were observed (Bailey et al., 1999):
Some of the clones did not produce any signal. It was thought that these cDNAs represented genes with low abundance mRNAs that northern blotting was too insensitive to detect.

Some of the clones hybridised weakly to all of the RNA samples. These cDNAs represented constitutively expressed genes with low abundance mRNAs that had not been eliminated by the first round colony screening. It seems likely that these colonies failed to hybridise to all three probes because the bacteria on the original plate had not transferred evenly to all three replica filters.

Two of the cDNAs (D13/3D and P8/8A) appeared to have a novel pattern of expression. They were not expressed in amoebae but were expressed in developing cells and plasmodia. They were expressed at peak levels in the developing cell culture. The expression of these cDNAs was analysed more extensively (section 1.3.3.3).

The fact that none of the 12 cDNA clones represented amoebal-specific or plasmodium-specific genes suggests the subtraction procedure worked.

1.3.3.3 The expression of D13/3D and P8/8A during the amoeba-plasmodium transition
To further characterise the expression pattern of D13/3D and P8/8A, a more detailed analysis was carried out by northern blotting with more RNA samples. This analysis is published in Bailey et al. (1999). In order to confirm that D13/3D and P8/8A had a novel expression pattern, their expression was compared to that of representative constitutively expressed, amoeba-specific and plasmodium-specific genes (Bailey et al., 1999). Actin was used as a positive control, to check that the blots contained intact RNA, because it is constitutively expressed (Hamelin et al., 1988) and its mRNA is present at approximately equivalent levels in amoebae and plasmodia (Pallotta et al., 1986). Actin was always the last probe to be hybridised to a blot; as it constitutes such a large proportion of the total mRNA in amoebae and plasmodia (approximately 4-5%; Pallotta et al., 1986), it is hard to strip from the blot. Two actin loci, ardB and ardC are responsible for generating the vast majority of actin in amoebae and plasmodia (Hamelin et al., 1988). The northern blots were probed with a cDNA probe, pPpA35 which hybridises to the ardB and the ardC transcripts; both are approximately 1400nt in length (Pallotta et al., 1986; Hamelin et al., 1988).

Three northerns blots were made which contained 10μg of total RNA isolated from (i) CL amoebae (ii) LU352 amoebae (iii) CL microplasmodia (iv) CL macroplasmodia and (v) a variety of CL developing cell cultures, each of which contained a different proportion of cells committed to apogamic development, ranging from 1-56% (Table 2.1; section 2.2.2; Bailey et al., 1999). Each northern blot was probed at high stringency (65°C) with at least two probes in separate hybridisation experiments. The first northern blot was probed for
the amoeba-specific and the plasmodium-specific profilins, proA and proP (Binette et al., 1990), and constitutively expressed actin (pPpA35: Hamelin et al., 1988; Fig. 1.3). The second northern blot was probed for D13/3D (redA) and actin (Fig. 1.3). And the third northern blot was probed for P8/8A (redB) and actin (Fig. 1.3). The actin signals from two of these blots are shown in Fig. 1.3; the actin signal in the top panel was derived from the blot that was probed for proA and proP, and the actin signal in the bottom panel was derived from the blot that was probed for D13/3D (redA) and is representative of the actin signal obtained on the blot that was probed for P8/8A (redB) (Bailey et al., 1999). The probing was repeated several times and the results shown in Fig. 1.3 are representative of those obtained (Bailey et al., 1999).

As expected, the actin probe hybridised to a mRNA of 1400nt (Pallotta et al., 1986). All of the RNA samples contain easily detectable intact actin transcript (Fig. 1.3, top and bottom panels). When looking at the actin signal on the blot in the bottom panel it is clear that the A (CL) RNA sample contains less actin transcript than any of the other samples on the blot. It is thought that this sample transferred badly (J. Bailey pers. com). The amoebal profilin, proA hybridised to two transcripts of 500 and 600 nucleotides (nt) in length, while the plasmodium-specific profilin, proP hybridised to a 520nt transcript (Fig. 3.1); these data are in agreement with previous findings (Binette et al., 1990). The two proA transcripts are generated from a single gene, but have different amounts of untranslated sequence (Binette et al., 1990). The proA mRNA is present in the RNA from amoebae and the developing cell cultures, but is absent in the RNA from microplasmodia and macroplasmodia (Fig. 1.3). Thus, as expected, proA is expressed in amoebae but not plasmodia (Binette et al., 1990). The proP mRNA is absent in the RNA from amoebae and the developing cell culture in which 1% of the cells were committed to development (Fig. 1.3). The proP mRNA was detected in the RNA from all of the other developing cell cultures, and the amount of proP transcript increased as the percentage of committed cells increased (Fig. 1.3). The plasmodial RNA samples contained the most proP mRNA (Fig. 1.3). Thus, as expected, proP is expressed in plasmodia, but not in amoebae (Binette et al., 1990).

If proA was exclusively expressed in amoebae, its expression would decrease as the percentage of amoebae that were present in the developing cultures decreased, e.g the 1% culture, in which 99% of the cells were amoebae would contain approximately twice as much proA transcript as the 56% culture, in which just 44% of the cells were amoebae. As the proA bands are roughly the same intensity in all of the developing cell culture RNAs (Fig. 1.3), proA must be expressed in uninucleate developing cells and possibly in
binucleate cells and quadrinucleate cells, like other amoeba-specific genes, e.g. α3 tubulin (Solnica-Krezel et al., 1990 and 1991).

Solnica-Krezel et al. (1998) used northern blotting to examine the expression pattern of the plasmodium-specific betC gene, which encodes the plasmodium-specific β2 tubulin, throughout development. The authors detected betC mRNA in developing cultures in which just 2% of the cells were committed to development. Using anti-β2 tubulin antibodies, Solnica-Krezel et al. (1990 and 1991) showed that β2 tubulin is first detected in uninucleate developing cells. The expression pattern of betC, as determined by northern blotting, is identical to that of proP (Solnica-Krezel et al., 1988; Bailey et al., 1999). Thus, it seems likely that proP is first expressed in uninucleate developing cells, like other plasmodium-specific genes, e.g. betC (Solnica-Krezel et al., 1988, 1990 and 1991).

Once the expression patterns shown by constitutively expressed (actin), amoebal-specific (proA) and plasmodium-specific (proP) genes had been determined, the expression of D13/3D and P8/8A was analysed to see if these cDNAs really did represent a class of genes that have a novel pattern of expression. D13/3D (redA) and P8/8A (redB) both hybridised to 800nt transcripts (Fig. 1.3). The redA and the redB mRNA could not be detected in the amoebal RNA samples (Fig. 1.3). Trace amounts of the redA and redB mRNA were detected in the RNA from the culture in which 4% of the cells were developing. Easily visible redA and redB transcript was seen in the RNA from all of the other developing cell cultures, and their transcript levels increased with the percentage of committed cells (Fig. 1.3) to reach a maximum in the 39% and 56% samples. The redA and redB transcripts were detected at lower levels in the RNA from microplasmodia and only trace amounts of transcript could be detected in the RNA from macroplasmodia (Fig. 1.3).

The expression pattern shown by redA and redB is markedly different from that of actin and proA, because they are not expressed in amoebae (Fig. 3.1). Although the expression pattern of the plasmodium-specific, proP is similar to that of redA and redB in the developing cell cultures with a low percentage of committed cells, in cell cultures with a higher proportion of developing cells and in microplasmodia and macroplasmodia, their expression patterns clearly differ; proP expression peaks in microplasmodia and macroplasmodia, whereas redA and redB expression peaks in the developing cultures with 39% and 56% committed cells and declines thereafter (Fig. 3.1). As redA and redB do not share the same expression pattern as constitutively expressed, amoebal-specific or plasmodium-specific genes, they represent a new class of genes that are expressed at high
levels during development. These genes were named red genes (regulated in development) and D13/3D and P8/8A were renamed redA and redB respectively (Bailey et al., 1999).

The subsequent analysis of the redA and redB cDNAs showed that they represented single-copy genes (Bailey et al., 1999). The deduced protein sequence of redA contained no conserved domains or motifs and was not homologous to other known proteins (Bailey et al., 1999). In contrast, the deduced protein sequence of redB contained a pair of calcium binding domains and was significantly homologous to the sarcomere calcium binding proteins (SCPs) from shrimp and Drosophila melanogaster (Bailey et al., 1999). The first members of the SCP family were found in invertebrate muscle, where they were thought to act as calcium buffers (Wnuk et al., 1982). But more recently, SCPs have been found in nerve cells which suggests their role is more complex than first thought (Kelly et al., 1997). The exact function of these proteins remains unclear.

1.4 DNA TRANSFORMATION IN P. POLYCEPHALUM

Stable DNA transformation provides a way of altering the nuclear genome of an organism. It allows the function of a gene of interest to be examined by either removing the function of that gene or by mis-expressing the gene. Burland et al. (1992a) began the development of a transformation system in P. polycephalum by creating a transient transformation system. A stable transformation system was later developed.

1.4.1 Transient transformations

Reporter genes generally encode enzymes that catalyse reactions which generate a readily detectable product (reviewed in Alam and Cook, 1990). Reporter gene systems allow the indirect analysis of the expression of a gene of interest using three basic steps: (i) the promoter of the gene of interest is translationally fused to the reporter gene, (ii) the construct is transformed into cells, and (iii) the cells are assayed for the product of the reporter gene. Generally, the amount of reporter gene product that is detected is proportional to the transcriptional activity that the promoter confers (Alam and Cook, 1990). Thus, reporter gene systems indirectly allow the expression pattern and the expression levels of a gene of interest to be determined.

The early transformation experiments used the bacterial chloramphenicol acetyl transferase gene, cat, as a reporter. The cat product catalyses the transfer of acetyl groups from acetyl-coenzyme A to chloramphenicol, and Cat activity is determined by biochemical analyses (reviewed in Alam and Cook, 1990). Burland et al. (1992a) transformed axenically grown amoebae by electroporation with a plasmid which contained cat under the control of either the promoter of the constitutively expressed actinB gene, PardB, or the promoter of the constitutively expressed actinC gene PardC, and detected Cat activity in the transformant
cells. The authors varied the experimental procedures and the electrical parameters of the electroporation to optimise the transient transformation procedure. These studies showed that PardC is a stronger promoter than PardB, since amoebae that were transformed with PardB-cat had higher levels of Cat activity than cells that were transformed with PardC-cat. By placing cat under the control of various exogenous promoters, Burland et al. (1992a) found that amoebae only recognised P. polycephalum promoters.

Later transformation experiments used the firefly luciferase gene, luc, as a reporter (de Wet et al., 1985 and 1997). The luc product oxidises luciferin to produce light and Luc activity is measured by determining the light output of the cells using a luminometer. Bioluminescent systems have two main advantages over cat systems; Luc assays are 30-1000 times more sensitive that Cat assays depending on how the light output is measured (de Wet et al., 1985), and Luc assays are less expensive and less time consuming than Cat assays (de Wet et al., 1985; Alam and Cook, 1990; Wood, 1990). Bailey et al. (1994) transformed axenically grown amoebae by electroporation with plasmid which contained luc under the control of PardB or PardC. Again, the authors varied the experimental procedures and the electrical parameters of the electroporation to optimise the transient transformation procedure. In agreement with the work of Burland et al. (1993a) they found that PardC was a stronger promoter than PardB. The authors also tested a sample of transformed cells for Luc activity over a period of 45 hours (Bailey et al., 1994). Cells that were transformed with PardB-luc and PardC-luc gradually lost their Luc activity over this time, which showed the plasmids were not maintained in the cells. The rate at which Luc activity was lost was faster than could be explained by vector dilution resulting from cell doubling, which suggests the plasmid is degraded by nucleases (Bailey et al., 1994).

Before conducting transient or stable transformation experiments, the optimum electrical parameters for the electroporation of the axenically grown cell line are determined by transient transformations. This is necessary because amoebae change after prolonged growth in axenic medium (Bailey et al., 1994).

1.4.2 Stable transformations

Only one selectable marker has been successfully used in P. polycephalum; the bacterial hygromycin phosphotransferase gene (hph: Gritz and Davies, 1983). Cells that carry expressing copies of hph are resistant to the antibiotic, hygromycin (Hyg). Burland et al. (1993a) first achieved stable transformation in P. polycephalum. The authors transformed amoebae by electroporation with plasmid containing PardC-hph-TardC (TardC is the transcriptional terminator of the actinC gene) and obtained HygR transformants at a frequency of approximately 1x10^-8 per cell. Burland and Bailey (1995) and Burland et al.
(1992b) later varied the experimental procedures of the transient and stable transformations to optimise the protocols, and published reliable methods for transforming *P. polycephalum* amoebae.

Burland and Pallotta (1995) achieved homologous gene replacement in *P. polycephalum*. The authors targeted the actinD gene, *ardD*, because it is not expressed in amoebae and the amoebal growth of any knockout strains would be unaffected. The authors transformed axenically grown amoebae with various constructs containing a mutated copy of *ardD* and *PardC-hph* and obtained a total of 126 Hyg\(^R\) transformants. The most successful transformation experiment yielded transformants at a frequency of 5.1\(\times\)10\(^7\) per cell. The analysis of 38 of the transformants showed that approximately 5% of the integration events were homologous (Burland and Pallotta, 1995). When the transformation efficiencies of the various constructs were compared, it became clear that (i) linear DNA integrates into the genome more frequently than circular DNA, and (ii) smaller pieces of DNA integrate into the genome more frequently than larger pieces of DNA (Burland and Pallotta, 1995).

1.5 THE *P. POLYCEPHALUM* GENOME

The C value in *P. polycephalum* is 0.3pg/nucleus, which corresponds to a haploid genome of approximately 3\(\times\)10\(^8\) bp (Mohberg *et al.*, 1973; Mohberg, 1977). Thus, the *P. polycephalum* genome is approximately 20 times larger than the yeast genome and one tenth of the size of the human genome. Studies have shown that the haploid chromosome number of different strains ranges from 25-75 chromosomes per haploid nucleus, and averages approximately 40 (Mohberg *et al.*, 1973; Mohberg, 1977).

Reassociation kinetic analysis has shown that the nuclear DNA of *P. polycephalum* is comprised of three sequence components; 6% is foldback sequence (inverted complementary sequences), 31% is repetitive sequence and 63% single-copy sequence (Hardman *et al.*, 1980). Hardman *et al.* (1980) showed that the repetitive sequences are interspersed with single-copy sequence in a large portion of the *P. polycephalum* genome.

1.6 DNA REPLICATION

Eukaryotic chromosomes are replicated from multiple initiation sites that are dispersed along the length of the chromosome (Hand, 1978). The chromosomal origins of *Saccharomyces cerevisiae* have been characterised in detail. There are specific replication initiation sites within *S. cerevisiae* chromosomes that colocalise to genetic elements that confer origin function (reviewed in Clyne and Kelly, 1997; Gilbert, 1998). These genetic elements, called autonomously replicating sequences (ARS) are approximately 100-200bp in size and are composed of an 11bp consensus sequence element, the ARS consensus
sequence (ACS), and 2-3 motifs that are divergent in sequence but functionally important. Proteins that interact with *S. cerevisiae* origins have been identified (reviewed in Diffley, 1995; Gilbert, 1998). Basically, a complex of six proteins called the origin recognition complex (ORC) remains bound to the ACS and the adjacent B1 motif throughout the cell cycle. The ORC nucleates the assembly of a complex of proteins called the pre-replicative complex (pre-RC) during late mitosis. The pre-RC is converted to an active RC after a series of events, leading to the initiation of DNA replication (Diffley, 1995; Gilbert, 1998).

The situation is more complex in higher eukaryotes as there has been difficulty in defining specific initiation sites and the genetic elements that confer origin function (reviewed in Clyne and Kelly, 1997; Gilbert, 1998). Some studies have shown that the initiation of replication occurs at multiple sites within broad ‘initiation zones’. For example, a chromosomal origin that is used for the amplification of the chorion genes in *Drosophila* follicle cells was mapped to a region which contains an amplification control element (ACE: Heck and Spradling, 1990). A closer examination of the locus showed that the initiation of replication was occurring at a multiple initiation sites flanking the ACE. A similar situation was observed in the dihydrofolate reductase locus of Chinese hamster cells. The authors found that although replication initiation was confined to a 30kb region, the specific site of initiation varied randomly within this zone (Vaughn *et al*., 1990). Other studies showed that the initiation of replication can occur at highly localised sites (reviewed in Clyne and Kelly, 1997; Gilbert, 1998). For example, the human β-globin (Aladjem *et al*., 1995) and the lamin B2 loci (Kumar *et al*., 1996) contain origins that can be mapped to a region of a few kilobases or less. The genetic elements that confer origin function in higher eukaryotes have not been identified.

There is a positive correlation between the transcriptional activity and the replication timing of a gene; active genes are generally early replicating (Hatton *et al*., 1988). Studies by Dhar *et al.* (1988) and Epner *et al.* (1988) showed that genes that were highly expressed in one cell-type, were replicated in early S-phase. Yet in cell types where they were inactive, the same genes were replicated later in S-phase. Transcriptional promoters that also function as origins of replication have been identified in several genes, including the human heat shock protein 70 (Taira *et al*., 1994) and the c-MYC genes (Vassilev and Johnson, 1990).

CpG islands are stretches of non-methylated DNA that are rich in the dinucleotide CpG (reviewed in Antequera and Bird, 1999). CpG islands are associated with the promoters of approximately 50% of mammalian genes, including all of the housekeeping genes and approximately half of the tissue specific genes (Antequera and Bird, 1999). The existence
of non-methylated CpG islands has long been a subject of interest. The discovery that some transcriptional promoters function as origins of DNA replication prompted Antequera and Bird, (1999) to propose that CpG islands are the genomic footprints of promoters that are also origins of DNA replication. In this model, promoters that are active during early embryogenesis recruit proteins that are able to initiate DNA replication. The binding of these proteins generates methylation free footprints that are transmitted through somatic cell divisions by maintenance methylation. While, promoters that are inactive during early development do not recruit replication complexes and any CpG dinucleotides are not protected from de novo methylation. The model predicts that half of the tissue specific genes are expressed during embryogenesis, since this proportion of tissue specific genes are associated with CpG islands. The authors suggest that after cell specialisation the CpG island associated genes that were active during embryogenesis may become transcriptionally repressed and late replicating, while the tissue specific genes that were inactive during embryogenesis may become transcriptionally active and early replicating. If this model is correct, it explains the correlation between the early replication and transcriptional activity of a gene.

1.6.1 DNA replication in \textit{P. polycephalum}

The plasmodium is ideal for DNA replication studies since a plate grown plasmodium contains up to $10^8$ nuclei with naturally synchronous mitotic cycles. Braun \textit{et al.} (1965) showed that there is a defined temporal order of replication in the plasmodial S-phase by labelling replicons in two consecutive S-phases. The experiments used BromodeoxyUridine (BrdU), an analogue of thymidine that has an increased molecular weight. The late replicating DNA was labelled with $[^3]$H thymidine in the first S-phase, and the early replicating DNA was labelled with BrdU in the second S-phase. When the DNA was extracted from the plasmodium, the tritiated DNA fraction did not coincide with the density-shifted, BrdU labelled DNA fraction. Thus, the DNA that was replicated during the first half of S-phase was also replicated during the first half of the consecutive S-phase. This work was confirmed and extended in a later study by Braun and Wili, (1969) which showed that sub-fractions of the \textit{P. polycephalum} genome (each representing approximately one fifth of the genome) were replicated at the same time during consecutive S-phases.

Studies involving incorporation of radioactive precursors into replicating DNA and flow cytometry experiments showed that the rate of DNA replication is not constant throughout the 3 hour S-phase (Braun \textit{et al.}, 1965; Kubbies and Pierron, 1983). During the first 90 min of S-phase, DNA synthesis occurs at a rate of approximately 5000kb/min/nucleus. This drops to a rate of approximately 1200kb/min/nucleus for the remaining 90 min of S-
phase. Approximately 75% of the genome is replicated during the first half of S-phase. Experiments conducted by Funderud et al. (1978b) showed that the replication within a replicon is bidirectional and that both of the replication forks move at the same rate. In these experiments, a plasmodium was given a pulse of BrdU at the onset of S-phase. After a chase, the DNA was exposed to UV irradiation, which breaks the BrdU-substituted DNA. Alkaline sucrose gradient analysis showed that the newly synthesised DNA had been cut into two strands of equal length by the UV, which showed the BrdU-substituted origins were located in the centre of the replicons.

Alkaline sucrose gradient analysis, electron microscopy and DNA fibre autoradiography have shown that \textit{P. polycephalum} replicons range from 12-60kb in size and average approximately 35kb. Each replicon is extended at a rate of approximately 1.2kb/min (Funderaud et al., 1978a and 1979). Bénard and Pierron (1990) examined the newly synthesised, or nascent strands of an individual replicon. By looking at the size of the nascent strands at different timepoints in S-phase they established that the replicon was elongating at a rate of approximately 1kb/min, in agreement with the findings described above. In this study, the size difference between the smallest and the largest nascent strand at any one time was approximately 5kb, which indicates a 5 min asynchrony between nuclei, within the approximately 10 hour mitotic cycle.

In order to investigate the relationship between the expression and replication timing of \textit{P. polycephalum} genes, 18 genes were examined in several independent investigations. The five genes that were not expressed in plasmodia (Pallotta et al., 1986; Pierron et al., 1989; Hamelin et al., 1988) were replicated at different times during S-phase; one in early S-phase (LAV5-1: Pierron et al., 1989), two in mid S-phase (LAV3-1 \textit{[proA]} and LAV3-3: Pierron et al., 1989) and two in late S-phase (LAV2-1: Pierron et al., 1989; \textit{ardA}: Pierron et al., 1984). While, most of the genes that were expressed in plasmodia (Pallotta et al., 1986; Pierron et al., 1989; Hamelin et al., 1988; Carrino et al., 1987) were replicated in early S-phase; of 13, ten were replicated in early S-phase (LAV1-1 \textit{[hapP]}, LAV1-2 \textit{[plasmin B]}, LAV1-3 \textit{[plasmin C]} and LAV1-5 \textit{[proP]}: Pierron et al., 1989; \textit{altB}: Cunningham and Dove, 1993; \textit{ardB}, \textit{ardC} and \textit{ardD}: Pierron et al., 1984; H4-1 and H4-2: Bénard and Pierron, 1999), one was replicated in mid S-phase (\textit{altA}: Cunningham and Dove, 1993) and two were replicated in late S-phase (LAV1-4 \textit{[php]} and LAV3-2: Pierron et al., 1989). It is clear that genes that are active in the plasmodium are generally replicated during early S-phase. Only one of the 13 active genes was replicated in late S-phase. In contrast, most of the inactive genes were replicated in mid or late S-phase. Only one of the five inactive genes was replicated in early S-phase. The results indicate that although active genes have a tendency to replicate early, and inactive genes have a tendency to
replicate late, the replication of both sets of genes is not restricted to any compartment of S-phase.

An earlier study by Pierron et al. (1982) provided evidence for a coupling of replication and transcription in *P. polycephalum*. The authors looked at electron microscope spreads of nuclear chromatin from a plasmodium in early S-phase and observed nascent transcripts in the centre of many early replicons. These data suggested a coincidence between transcriptional promoters and replication initiation sites, because origins of replication are also located in the centre of replicons. In order to further investigate the relationship between transcription and replication origins in *P. polycephalum*, the replication initiation sites that are located close to six early replicating genes were mapped in several independent studies using two-dimensional (2D) gel electrophoresis and sometimes alkaline gel electrophoresis (Bénard and Pierron, 1992; Diller and Sauer, 1993; Bénard et al., 1996; Bénard and Pierron, 1999).

All of the six genes are abundantly transcribed in the plasmodium. Two are cell-type specific; the plasmodium-specific profilin gene, *proP*, (LAV1-5: Pallotta et al., 1986; Pierron et al., 1989) and the plasmodium-specific gene, LAV1-2 (Pallotta et al., 1986; Pierron et al., 1989). And four are constitutively expressed; the actin genes *ardB* and *ardC* (Hamelin et al., 1988) and the histone 4 genes H4-1 and H4-2 (Carrino et al., 1987). The studies showed that the promoter regions of *proP*, *ardB*, *ardC*, H4-1 and H4-2 contain replication origins that are activated at the onset of S-phase in the plasmodium (Bénard and Pierron, 1992; Diller and Sauer, 1993; Bénard et al., 1996; Bénard and Pierron, 1999). In contrast, LAV1-2 is situated in the middle of two ‘weak’ replication origins that are not activated 100% of the time. When the origins are activated, it is at the onset of S-phase. The origins were approximately 10kb apart and when they are both activated on the same chromatin strand, they generate replication forks that converge towards LAV1-2 and clash within the gene region (Diller and Sauer, 1993). This data confirms that there at there is a coupling of replication and transcription in *P. polycephalum*, in agreement with the electron microscope observations.

Pierron et al. (1999) provided the first direct evidence of a coincidence between a transcriptional promoter and a replication origin in *P. polycephalum*. It has been established that the 1.1kb fragment located directly upstream of the actin gene, *ardC*, is a transcriptional promoter (Burland et al., 1992a; Bailey et al., 1994). As discussed above, a replication initiation site has also been mapped to this region (Bénard et al., 1996). Pierron et al. (1999) translationally fused *PardC* to the selectable marker, *hph*, transformed the linear construct into amoebae and selected stable transformants on the basis of Hyg resistance. The analysis of one of the transformants showed that the presence of the
construct had converted a late replicating region of the genome into an early replicating
one. Thus, the 1.1kb fragment *PardC* is both a promoter of RNA transcription and an
origin of DNA replication.

1.7 AIMS

*P. polycephalum* amoebae and plasmodia are very different cell types that differ in gene
expression, cellular organisation and behaviour. Genetic and immunological studies have
shown that the changes in gene expression and cellular reorganisation that result in the
formation of a plasmodium are initiated during the extended cell cycle of the APT
(Sweeney *et al*., 1987; Solnica-Krezel *et al*., 1988, 1990 and 1991). The study of non-
selfing mutant strains led to the identification of several mutations that block plasmodium
development, but do not affect amoebal growth (Wheals, 1973; Anderson and Dee, 1977;
Solnica-Krezel *et al*., 1995). Although most of these *npf* (no plasmodium formation)
mutations are linked to *matA*, some are not. The *matA*-unlinked *npf* mutations were studied
extensively (Bailey *et al*., 1992a and b; Solnica-Krezel *et al*., 1995). The observation that
some aspects of plasmodial development occurred normally in the mutant strains led to the
idea that plasmodium development is controlled by several independent regulatory
pathways that collectively result in the formation of a plasmodium (Bailey *et al*., 1992a;
Burland *et al*., 1993a; Solnica-Krezel *et al*., 1995). In order to identify development
specific genes, Bailey *et al.* (1992a) constructed a cDNA library and enriched the library
by subtraction for genes that were expressed during development. The screening of this
library led to the identification of a novel class of genes that are expressed primarily during
development (Bailey *et al*., 1999). These genes were called *red* (regulated in
development) genes and their function is unknown (Bailey *et al*., 1999). My initial aim
was to isolate more *red* genes from the subtracted cDNA library. My second aim was to
characterise any isolated gene(s), since the analysis of the *red* genes may lead to an
understanding of how plasmodium development is controlled. And my third aim was to
create a luciferase reporter gene system that could be used in stable transformations to
examine the activity of the *red* gene promoters, and the function of the *red* genes by mis-
expression.

DNA replication studies in the plasmodium have shown that the replication of active genes
is generally confined to early S-phase, while the replication of inactive genes is not
confined to any compartment of S-phase (Pierron *et al*., 1989). Subsequent studies showed
that genes that are abundantly transcribed in the plasmodium generally contain replication
initiation sites within their upstream promoters, that fire at the onset of S-phase (Bénard
The *red* genes were of interest for replication studies because they are expressed at low
levels in the plasmodium, and genes that within this category have not previously been
studied. My final aim was to examine the replication timing of the red genes in the plasmodium, and to map the replication initiation sites that are close to any early replicating red genes.
Figure 1.1: The heterothallic life-cycle of *Physarum polycephalum*
Figure 1.2: The timing of events in the amoebal-plasmodial transition in the apogamic strain, CL (Bailey et al., 1987)

Amoebal cell cycle
1.0 units

Open mitosis

Loss of ability to flagellate

Closed mitosis

Synchronous closed mitoses

Haploid amoeba

Commitment

Extended cell cycle
2.3 units

Shortended cell cycle
0.7 units

Haploid plasmodium
Total RNA was isolated from LU352 amoebae (A[LU352]), CL amoebae (A[CL]), CL microplasmodia and macroplasmodia (Mi; Ma) and from a variety of developing CL cultures (1-56%; Table 2.1) as described in section 2.2.2. 10μg of each RNA sample, and 3μg of RNA size markers (Promega) were denatured and run on a 1.1% agarose gel as described in section 2.4.2. The RNA was transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.2).

The cDNA clones for actin (Hamelin et al., 1988), proA and proP (Binette et al., 1990) and redA and redB (D13/3D and P8/8A respectively; Bailey et al., 1999) were radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised at high stringency (65°C) to the northern blots. The blots were exposed at -80°C with an intensifying screen. Actin hybridised to a 1400nt mRNA, proA to 600nt and 500nt mRNAs, proP to a 520nt mRNA, redA to an 800nt mRNA and redB to an 800nt mRNA.
CHAPTER 2: MATERIALS AND METHODS

2.1 P. POLYCEPHALUM CULTURE AND TRANSFORMATION

2.1.1 Gene loci and strains

Gene loci

- **matA**: key mating-type locus affecting the initiation of development.
- **gadA**: greater asexual differentiation. A temperature sensitive mutation linked to **matA** which allows apogamic development at 26°C. Development is blocked at 30°C.
- **npfC5**: no plasmodium formation mutation. A mutation linked to **matA** and **gadA** which blocks apogamic development. Reversion to wild type occurs at high frequency.
- **matB, matC**: mating-type loci which affect the frequency of amoebal fusion.
- **fusA, fusC**: Plasmodial fusion loci.

Strains

**P. polycephalum**

- LU352: **matA**2 **gadA**h **npfC**5, **matB**3, **matC**1, **fusA**1, **fusC**1.
- CL: **matA**2 **gadA**h **npfC**+, **matB**1, **matC**1, **fusA**2, **fusC**1.

**Escherichia coli**

- XL1-Blue: **recA**1 **endA**1 **gyrA**96 **thi**-1 **hsdR**17 **supE**44 **relA**1 **lac** [**F’ proAB lacI**³ **ZAM15 Tn10 (Tet’)].

2.1.2 Media

All solutions were made in water unless stated otherwise.

- **SDM**: Semi-defined medium (Blindt et al., 1986).
- **SDM-agar**: 1.5% w/v agar in 50% v/v **SDM** (Blindt et al., 1986).
- **DSDM**: Dilute **SDM** (Blindt et al., 1986). 1.5% w/v agar in 6.25% v/v **SDM**.
- **DSDM-hyg**: DSDM containing hygromycin B at a concentration of 100µg ml⁻¹.
- **DSPB**: Dilute **SDM** with phosphate buffer (Dee et al., 1997). **DSDM** containing 10mM sodium phosphate buffer pH 6.8 (Sambrook et al., 1989).
- **SDM+PSA**: **SDM** containing penicillin, streptomycin and ampicillin at a concentration of 150µg ml⁻¹, 250µg ml⁻¹ and 100µg ml⁻¹ respectively.
- **LB broth**: Luria Bertani broth (Sambrook et al., 1989). 0.5% w/v bacto-peatone and 0.25% w/v bacto-yeast in 0.25M NaCl pH 7.0.
**LB-agar**  
LB broth containing 1.5% w/v agar (Sambrook et al., 1989).

**LB-amp**  
LB-agar containing ampicillin at a concentration of 100 μg ml⁻¹.

**SOC broth**  
2% w/v bacto-tryptone, 0.5% w/v bacto-yeast, 8.5 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ and 20 mM glucose (Sambrook et al., 1989).

**SBS**  
Live standard bacterial suspension (Burland et al., 1981) prepared by washing a lawn of *E. coli* (strain B145; Leicester University) off an LB-agar plate with 5 ml of sterile distilled water. SBS is used as an amoebal food source.

**FKB**  
Formalin-killed bacteria (Dee, 1986; Dee et al., 1997). FKB was sometimes used as the amoebal food source instead of SBS. *P. polycephalum* transformants are selected on the basis of hygromycin (hyg) resistance. FKB was used as the amoebal food source during transformant selection because live bacteria could potentially affect amoebal growth or hyg degradation.

### 2.1.3 Excysting and encysting amoebae

Amoebae naturally form cysts when they are cold or when starving, and this is an ideal way to store them. Cysts were stored at -80°C in 15% v/v glycerol (Dee, 1986). To grow active amoebae, 20 μl of the stock was mixed with 100 μl of SBS or FKB, and the suspension was spread onto a 9 cm DSPB plate with a glass spreader. The plate was then incubated at 26°C for 1-2 days to allow the cells to excyst and begin to grow. Once the cells were growing, the plate was transferred to 30°C to block plasmodium development, but allow amoebal growth to continue. The amoebae were sub-cultured when the cells became confluent. In this procedure 100 μl of SBS or FKB was placed into the middle of a fresh DSPB plate. Some of the amoebae were then transferred from the old plate to the bacterial suspension on the new plate with a glass spreader and plate was incubated as before. Plates containing encysted cells were sometimes stored at 4°C, since cysts remain viable at this temperature for several weeks.

To prepare frozen stocks, amoebae on a DSPB plate were incubated at 4°C to induce excystment. The cysts were later washed off the plate with 3 ml of ice cold 15% v/v glycerol and the suspension was aliquoted into several cryo-tubes and stored at -80°C.

### 2.1.4 Axenic culture of amoebae

To ensure the cells grew well in axenic culture, plate-grown amoebae were encysted before they were transferred to the axenic media. In this procedure, amoebae were washed off a DSPB plate with 2 ml of flagellation buffer (10 mM phosphate buffer pH 6: Sambrook *et al.*
al., 1989). The cells were pelleted for 1.5 min at 400xg and the supernatant was removed. The cells were then resuspended in 3ml of the flagellation buffer and pelleted as before. This step removed a lot of the bacteria from the suspension, as they do not pellet in these conditions. It was necessary to remove the excess bacteria since their presence can hinder flagellation. The supernatant was removed and the cells were resuspended in 2ml of flagellation buffer. The number of cells per ml was calculated by counting the cells with a haemocytometer. The cells were then incubated at room temperature for several hours on a reciprocating shaker, to allow amoebal excystment and flagellation. The cells were counted with a haemocytometer at regular intervals to determine the proportion of flagellates. When approximately 50% of the amoebae had flagellated, the cells were pelleted as above and the supernatant was removed. The cells were then resuspended in a variable volume of SDM+PSA so that the cell density was 1-5x10^5 ml^-1. The cell suspension was then divided into 1.5ml aliquots and several tubes were incubated at 30°C on a shaker. The cells were sub-cultured to fresh SDM when the cell density reached approximately 1x10^7 ml^-1, to keep the cells in log-phase growth. When LU352 amoebae had been sub-cultured 15-20 times, they were transferred to 100ml of SDM in a conical flask with a screw cap lid, because the transformations required a large number of cells. The amoebae were then sub-cultured as before.

2.1.5 Amoebal transformations

Axenically grown LU352 transformed by electroporation as described in Bailey et al. (1994). In this procedure, the transforming DNA was added to an aliquot of 5x10^7 cells. The cell suspension was then transferred to a 4mm cuvette and the electroporated using a Bio-rad gene pulsar at the optimum transformation settings (see below). After electroporation, 4.2ml of SDM was added to each sample and the cells were given a 2.5 hour period of outgrowth at 30°C on a shaker.

2.1.5.1 Transient transformations

All of the transient transformation vectors were circular plasmids that carried the firefly luciferase gene, luc (de Wet et al., 1985; 1987). The cells were transformed with 1μg of a 6.6kb plasmid and the weight was adjusted for bigger plasmids so that the cells were always transformed with approximately the same number of plasmid molecules, e.g 1.43μg of a 9.6kb plasmid was transformed. After the 2.5 hour period of outgrowth, the number of cells per ml was determined by counting the cells with a haemocytometer, and the percentage cell survival was determined by comparing the total number of cells before and after the transformation. Approximately 1x10^7 cells were then removed from each sample and assayed for Luc activity using a luminometer, as described in Bailey et al. (1994). The luminometer measures light emission in relative light units (RLU) over a
period of 30 seconds. An empty tube was also assayed to determine the background light levels. This background activity reading was subtracted from the other readings before they were analysed. An aliquot of cells with no added DNA was subjected to the electroporation procedure to determine the background Luc activity of *P. polycephalum* amoebae. This reading was always the same as the empty tube reading because amoebae do not have inherent Luc activity.

Every 6-12 weeks the optimum transformation parameters for an LU352 axenic culture cell-line were determined (Bailey *et al*., 1994). This was necessary because amoebae change after prolonged culture in axenic medium. In this procedure, 1μg of *pPardC-luc* plasmid was added to 12 aliquots of 5x10^7 cells. The first sample was electroporated with capacitance, voltage and resistance settings of 2.5μF, 0.65kV and 400Ω respectively. The settings were changed for the 11 of the remaining aliquots; voltage settings of 0.65, 0.7, 0.75 and 0.8kV were combined with each of the following resistance settings, 400, 800 and 8 Ω. These settings were chosen because they are in the range of previously observed optimum parameters (Bailey *et al*., 1994). After electroporation, the cells were given a period of outgrowth and assayed for luciferase activity, as described above. The optimum electrical parameters were those that resulted in the highest light emission, and allowed at least 50% cell survival. These parameters were used in all subsequent transformation experiments.

2.1.5.2 Stable transformations
All of the stable transformation vectors were linear constructs that contained the hygromycin phosphotransferase gene, *hph* (Gritz and Davies, 1983). The cells were transformed with approximately 3μg of each vector. The transformation experiments included three controls; one aliquot of cells was transformed with 1μg of circular *pPardB-luc* as a positive control, and two aliquots of cells were transformed with no DNA as negative controls. After the 2.5 hour period of outgrowth the cell, the number of cells per ml was determined by counting the cells with a haemocytometer, and the percentage cell survival was determined by comparing the total number of cells before and after the transformation. Approximately 1x10^7 cells were then removed from the positive and one negative control and assayed for Luc activity to check that the transformation had worked. The other samples were transferred to 50ml of SDM+PSA in a conical flask with a screw cap lid and incubated at 30°C for 2 days on a reciprocating shaker. The transformant selection procedure was carried out as described in Burland and Bailey, (1995).

After 2 days the total number cells in each sample was determined by counting the cells with a haemocytometer. The cells were pelleted at 2500xg for 3 min, the supernatant was
removed and the cells were resuspended in 600-800μl of FKB. Equal volumes of the cell suspension were then plated onto 4-6 DSDM-hyg plates at a maximum density of 7.5x10^7 cells per plate. To check the plating efficiency, a sample of cells from the remaining negative control were diluted in SDM and plated onto two DSDM plates with 100μl of FKB at a density of 100 cells per plate. All of the plates were incubated at 26°C for 1-2 days. When the cells on the DSDM plates started growing, the plates were transferred to 30°C. After 10 days of incubation the number of amoebal colonies on the 2 DSDM plates was counted, and the plates were discarded. These plates gave an indication of what proportion of the cells survived the re-plating procedure. The selective plates were examined every few days for signs of colony growth and were finally discarded after 3-4 weeks of incubation.

Any colonies that grew on the selective plates were sub-cultured onto a fresh selective plate to confirm the designation of hyg resistance. In this procedure, 100μl of FKB was placed into the middle of a fresh DSDM-hyg plate and some of the amoebae were transferred from the colony on the old plate to the bacterial suspension on the new plate using a sterile toothpick. The plate was then incubated at 26°C for 1-2 days before transfer to 30°C to block plasmodium development. When the cells were growing well, the amoebae were sub-cultured onto two DSPB plates with FKB, as before. The cells on the first plate were used to make a frozen stock of the transformant (section 2.1.3). The cells on the second plate were grown to confluence and transferred to axenic culture (section 2.1.4). Axenically grown cultures of amoebae were used for DNA (section 2.2.3) and RNA (section 2.2.2) isolations.

2.1.6 Inducing plasmodium development
Amoebae, grown on a DSPB plate, were sub-cultured onto a DSDM plate with SBS as previously described (section 2.1.3), and the plate was incubated at 26°C. DSDM plates are more suited to plasmodial growth than DSPB plates because they have a lower pH. When the plasmodium was 0.5-1cm in diameter, the block of agar containing the plasmodium was excised from the plate, transferred to an SDM-agar plate and incubated at 26°C. The plasmodium was sub-cultured every 3-4 days when the cell had grown to cover the plate. In this procedure, a 1cm^2 block of agar containing the growing edge of the macroplasmodium was excised from the plate, transferred to a fresh SDM-agar plate and incubated at 26°C.

2.1.7 Axenic culture of plasmodia
A 1cm^2 block of agar containing the growing edge of a macroplasmodium was excised from the SDM agar plate, transferred to 50ml of SDM in a 500ml conical flask and
incubated at 30°C on a shaker (Daniel and Rusch, 1961). The microplasmodia were sub-cultured every few days to keep the cells in log-phase growth. In this procedure, approximately 2ml of the microplasmodia suspension was transferred to 50ml or 100ml of fresh SDM in a 500ml conical flask and incubated at 30°C as before. Axenically grown cultures of microplasmodia were used for DNA (section 2.2.4) and RNA (section 2.2.2) isolations.

2.2 NUCLEIC ACID ISOLATION

2.2.1 Quantification of nucleic acids by spectrophotometry

The concentration of isolated RNA and DNA was determined using a spectrophotometer. In this procedure, 1μl of the RNA or DNA to be quantified was diluted in 0.5ml of sterile distilled water and transferred to a 1ml quartz cuvette. Firstly, the absorbance of 0.5ml sterile distilled water was measured at wavelengths 260nm and 280nm with a Hitachi U-2000 spectrophotometer. These readings were called 0, and the absorbance of the diluted RNA or DNA was then measured at the same wavelengths relative to 0. The concentration of DNA and RNA was determined using the following equation:

\[
\text{DNA concentration (µg µl}^{-1}) = \frac{\text{OD}_{260} \times \text{dilution factor (500) x 50 or 40}}{1000}
\]

The OD_{260} and dilution factor was multiplied by 50 for DNA and by 40 for RNA. The purity of the isolated RNA or DNA was calculated by comparing the 260nm and 280nm absorbency readings. A pure DNA sample has a 260nm:280nm ratio of 1.8, while a pure RNA sample has a 260nm:280nm ratio of 2.0. As proteins absorb light at 280nm, impure samples have an increased 280nm absorbency reading, which lowers the 260nm:280nm ratio. If the 260nm:280nm ratios were significantly lower than expected, further phenol:chloroform:isoamyl alcohol extractions were carried out and the sample was re-examined.

2.2.2 RNA isolation

Total RNA was isolated from amoebae and microplasmodia and according to Chomczyński and Sacchi (1987), with modifications by Puissant and Houdebine (1990). The cells were harvested in different ways. Non-axenic amoebae were washed off plates with 5ml ice cold water. Axenic amoebae were pelleted by centrifugation at 400xg for 3 min and washed twice in 50ml of ice cold water. Microplasmodia were pelleted by centrifugation at 400xg for 3 min and washed twice in 50ml ice cold water. And macroplasmodia were transferred straight to the lysis solution. The cells were then lysed in a solution containing 4M guanidium thiocyanate, 25mM sodium citrate pH 7.0, 0.5% v/v sarcosyl and 100mM β-mercaptoethanol. Guanidium thiocyanate and sodium chloride are protein denaturants that lyse the cells and β-mercaptoethanol is an RNase inhibitor that
prevents the degradation of RNA. The RNA was separated from the DNA by phenol extraction under acidic conditions and several washes in 4M lithium chloride were carried out to remove contaminant polysaccharides from the sample (Puissant and Houdebine, 1990). This was an important step for plasmodia which contain lots of slime. The RNA was precipitated with Propan-2-ol, washed in 75% ethanol and pelleted by centrifugation. The pellet was then dried and resuspended in diethylpyrocarbonate (DEPC) treated water.

The RNA from the crossed microplasmodia (LU648xCH508) was isolated by Barber (1998) with a RNAzol kit (AMS Biotechnology) according to the manufacturers instructions, although the protocol was modified to include two lithium chloride washes (Barber, 1998). The RNA concentrations were determined using a spectrophotometer (section 2.2.1). To confirm that the isolated RNAs were intact, 3μg of each sample was denatured and run on a gel containing 1.1% w/v agarose in 1xMOPS (section 2.4.2). The RNA was stained with 100ng ml\(^{-1}\) ethidium bromide and visualised under UV illumination. If a smear of very small RNA molecules were seen, the RNA was thought to be degraded and the sample was discarded.

The RNA from the apogamic CL developing cultures was isolated by Bailey et al. (1999). To produce these cultures, Bailey et al. (1999) inoculated nine sets of 100 plates containing SBS with 5x10\(^5\) amoebae. The sets of plates were then incubated at 26°C for different times to allow plasmodial development to progress to different stages. At the appropriate time, the cells on one set of 100 plates were harvested for RNA isolation. In this procedure, the cells were washed off each plate with 5ml of ice cold water and the cell suspensions were pooled. A small volume of the pooled cell suspension was removed for analysis prior to RNA isolation so that the distribution of cell types that were present in the culture could be determined.

The proportion of amoebae and multinucleate cells in each sample were determined by counting the number of nuclei per cell in a sample of 500-1000 cells using a phase-contrast microscope. The proportion of committed cells that were present in each culture was determined by plating the dilutions of the suspension onto fresh DSDM plates and incubating the plates at 26°C for approximately 5 days. After 2-3 days the committed cells formed small plasmodia and after 4-5 days the non-committed cells formed amoebal colonies (Bailey et al., 1987). Since uninucleate developing cells could not be distinguished from amoebae by the microscopic analysis, the proportion of uninucleate developing cells present was calculated by subtracting the proportion of multinucleate cells from the proportion of committed cells. Table 2.1 shows the distribution of cell types that were present in the nine developing cultures from which RNA was isolated.
The RNA from the heterothallic developing cultures was isolated by Barber (1998). To produce these cultures, Barber (1998) mixed equal quantities of flagellates from each mating strain. Each mixture was then plated onto 100 9cm SM-2 plates (Blindt, 1987) as 60x10μl spots at a density of 3x10⁴ cells per spot. Each set of 100 plates were incubated at 26°C for different times to allow plasmodial development to progress to different stages. At the appropriate time, the cells on one set of 100 plates were harvested for RNA isolation, as above. The proportion of amoebae, fusion cells, zygotes and multinucleate cells was determined by microscopic analysis, as above. The proportion of committed cells that were present in each culture was determined by re-plating assays, as above. Barber (1998) found that the proportion of fusion cells, zygotes and multinucleate cells equalled the proportion of committed cells that were determined in the replating assays. Table 2.2 shows the distribution of cell types that were present in the six developing cultures from which RNA was isolated.

2.2.3 Isolation of genomic DNA from axenically grown amoebae

Genomic DNA was isolated from axenically grown amoebae using a modification of a method devised by D. Pallotta (Quebec). Although the DNA isolated in this way gave reasonable Southern blotting results it was not suitable for cloning or restriction enzyme mapping because it was not very pure. These experiments required cleaner genomic DNA, isolated using a different method (section 2.2.4).

Four tubes of amoebae, grown at 30°C to between 5x10⁶ and 1x10⁷ cells ml⁻¹, were pooled and pelleted at 400xg for 1.5 min. The pellet was washed in 5ml BSS (basal salt solution containing 14mM citric acid, 24mM KH₂PO₄, 4.3mM NaCl, 850μM MgSO₄ and 340μM CaCl₂ pH 5.0: Bailey et al., 1994), resuspended in the residual liquid and lysed by gentle pipetting in 1.5ml of 4M guanidium thiocyanate. Any cell debris was pelleted by centrifugation at 1400xg for 5 min and the supernatant was transferred to a fresh tube. The DNA was precipitated with 2 volumes of ethanol and a 0.1 volume of 3M sodium acetate pH 5.2, and pelleted at 1400xg for 3 min. The pellet was washed with 1ml of 75% v/v ethanol, transferred to a fresh eppendorf tube and re–pelleted by centrifugation at maximum speed in a microfuge for 1 minute. The supernatant was removed and the DNA pellet was slightly dried before being resuspended in 400μl of 1xTE (10mM Tris-HCl and 1mM EDTA pH 8).

The RNA was removed by treating the sample with DNase-free RNaseA at a concentration of 100μg ml⁻¹ for 15 min at 65°C. The DNA was extracted in phenol:chloroform:isoamyl
alcohol (50:49:1 v/v) for approximately 20 min. The sample was then centrifuged at 13800xg for 10 min and the aqueous phase was transferred to a fresh eppendorf tube. The DNA was precipitated with ethanol and sodium acetate and washed in 75% ethanol, as described above. The supernatant was discarded and the DNA pellet was dried and resuspended in 50μl of sterile distilled water. The DNA concentration was determined using a spectrophotometer (section 2.2.1).

2.2.4 Isolation of genomic DNA from microplasmodia
Genomic DNA was isolated from an axenically grown culture of CL or LU352 microplasmodia (section 2.1.7) as described in T'Jampens et al. (1997). This method yields pure DNA that is suitable for cloning and restriction enzyme mapping. The cells were gently homogenised in nuclear homogenisation solution (NHS: 0.25M sucrose, 0.1% w/v Triton X-100, 15mM CaCl₂ and 10mM Tris-HCl pH 7.5) to release the nuclei. Poorly homogenised material was removed by filtration through a milk filter and the remaining nuclei were washed once in NHS:percoll 7:2 v/v and twice in NHS. The percoll separates the nuclei from the slime by trapping the slime at the top of the tube. The supernatant was removed and the nuclei were lysed in nuclear lysis solution (NLS: 100mM EDTA, 0.5% w/v Triton X–100 and 25mM Tris-HCl pH 8). The suspension was then treated with 100μg ml⁻¹ Proteinase K and 1mg ml⁻¹ DNase-free RNaseA. The DNA was then cleaned by successive extractions with phenol, phenol:chloroform:isoamyl alcohol (50:49:1 v/v) and chloroform:isoamyl alcohol (49:1 v/v). The DNA was then precipitated with ethanol and sodium acetate and washed twice with 75% v/v ethanol. The supernatant was removed and the DNA pellet was dried slightly and resuspended in sterile distilled water. The DNA concentration was determined using a spectrophotometer (section 2.2.1).

2.2.5 Isolation of plasmid DNA
The following method produces pure plasmid DNA that can be used for cloning. Frozen stocks of bacterial strains were stored in 10% v/v glycerol at ~80°C. To start the cells growing, bacteria were transferred from the top of frozen stocks to an LB-amp plate with a wire loop and the plate was incubated at 37°C overnight. Individual colonies were then transferred from the plate to either 10ml or 100ml of LB-amp with a wire loop and the culture was incubated overnight on a shaker at 37°C. Plasmid DNA was isolated from the culture with a Qiagen Plasmid Midi or Maxi kit, according to the manufacturers instructions. The DNA concentration was determined using a spectrophotometer (section 2.2.1).
2.2.6 Rapid isolation of plasmid DNA

The following method produces plasmid DNA that is not very pure. Colonies of interest resulting from a bacterial transformation (section 2.3.5) were transferred from an LB-amp plate to 1ml of LB-amp in an eppendorf tube and to a fresh LB-amp plate using a wire loop. The plates and the broth were incubated at 37°C overnight. The next day, the plates were transferred to 4°C and plasmid DNA was isolated from the 1ml cultures using the 'lysis by boiling method' of Sambrook et al. (1989). In this procedure the cells were pelleted by centrifugation and lysed in STET (8% w/v sucrose, 5% v/v Triton X-100, 50mM EDTA and 50mM Tris pH8) containing 1μg of lysozyme. The cell debris was pelleted by centrifugation and the supernatant was transferred to a fresh eppendorf. The DNA was precipitated with propan-2-ol and pelleted by centrifugation. The pellet was then dried and resuspended in sterile distilled water. The plasmids were analysed by restriction enzyme digestion and gel electrophoresis. When a plasmid of interest was identified, the corresponding colony on the plate at 4°C was transferred to 10 or 100ml of LB-amp to enable a larger scale plasmid isolation to be carried out.

2.3 DNA CLONING

2.3.1 Restriction enzyme digestion

Restriction enzyme digests were carried out using Gibco BRL restriction enzymes, according to the manufacturers instructions.

2.3.1.1 Large scale plasmid digests

Plasmid DNA was digested for use in cloning or to obtain vector-free purified inserts that were used as probes. Between 5μg and 20μg of plasmid DNA (section 2.2.5) was digested with an excess of enzyme (2-3 units μg⁻¹) at 37°C overnight in a final volume of 100-200μl. After digestion, approximately 500ng of the digested DNA was run on a 0.8% agarose gel to check that the digest had gone to completion. If the digestion was incomplete more enzyme was added and the digest was incubated at 37°C for a few more hours. If the digestion was complete the DNA was extracted with one volume of phenol:chloroform:isoamyl alcohol 50:49:1 v/v and centrifuged at maximum speed in a microfuge for 10 min. The aqueous phase was then transferred to a fresh eppendorf and the DNA was precipitated with 2 volumes of ethanol and 0.1 volumes of 3M sodium acetate pH 5.2. The DNA was then incubated at –20°C for 45 min and pelleted at maximum speed in a microfuge for 15 min. The supernatant was removed and the pellet was washed in 75% v/v ethanol. After re-pelleting, the supernatant was removed and the pellet was dried and resuspended in sterile distilled water.
2.3.1.2 Small scale plasmid digests
Plasmid DNA, obtained by the 'lysis by boiling method' of Sambrook et al. (1989) (section 2.2.6) was usually in a 20μl volume. Approximately 0.5μl of this plasmid DNA, or 0.5μg of plasmid DNA extracted with a Qiagen kit (section 2.2.5) was digested with 1-2 units of enzyme at 37°C for 45 min in a final volume of 20μl. The digests were then loaded straight onto an agarose gel.

2.3.1.3 Genomic DNA digests
For Southern blots, approximately 3μg of genomic DNA from microplasmodia (section 2.2.4) was digested with an excess of enzyme (5-10 units μg⁻¹) at 37°C overnight in a final volume of 100-200μl. The DNA was then extracted with phenol:chloroform:isoamyl alcohol, ethanol precipitated, washed and resuspended in sterile distilled water, as described above.

For the genomic library, approximately 60μg of genomic DNA from microplasmodia was digested with an excess of XbaI and HindIII (5 units μg⁻¹) at 37°C overnight in a final volume of 300μl. The digested DNA was then loaded straight onto an agarose gel and restriction fragments of 1-2kb were gel isolated (section 2.3.3).

2.3.2 Agarose gel electrophoresis
Gels containing 0.8-1.5% agarose v/v in 1xTBE (90mM Tris-Cl, 90mM boric acid, 2mM EDTA pH 8: Sambrook et al., 1989) were prepared and electrophoresed in 1xTBE. A 0.1 volume of loading buffer was added to each of the DNA samples immediately prior to loading (0.25% w/v bromophenol blue in 40% w/v sucrose: Sambrook et al. 1989). Generally λ HindIII and/or φX174 HaeIII DNA size markers (Gibco BRL) were included on the gels. Gels were usually stained with 100ng ml⁻¹ ethidium bromide and photographed under UV illumination.

2.3.3 Isolating restriction fragments of interest
Restriction fragments were isolated for cloning or for use as probes and during the size fractionated genomic DNA library construction. Digested plasmid DNA was run on an agarose gel and stained with ethidium bromide at a concentration of 100ng ml⁻¹. The restriction fragment of interest was visualised using a hand held UV lamp and isolated from the gel using Na45-DEAE cellulose membrane, as described in Sambrook et al. (1989) or a QIAquick gel extraction kit (Qiagen), according to the manufacturers instructions. The isolated restriction fragments were resuspended in sterile distilled water. A small amount of the purified insert was usually run on an agarose gel with λ HindIII
and/or $\Phi$X174 HaeIII DNA size markers of known concentration. The gel was then stained with ethidium bromide, photographed under UV illumination (section 2.3.2) and the DNA concentration was estimated by comparing the intensity of the insert band to the intensity of several DNA size marker bands.

2.3.4 Vector and insert preparation
Most of the restriction fragments were cloned into pBluescript II (Stratagene) or pGem (Promega) which confer ampicillin resistance. The vector was usually digested with restriction enzymes that generated compatible ends for the cloning of a specific restriction fragment. When the vector did not contain suitable restriction sites it was digested with an enzyme that generated blunt ends, such as EcoRV. The ends of the insert were then blunted with the DNA polymerase I large (Klenow) fragment (Gibco BRL) before ligation. Klenow fills 5' overhangs with deoxynucleotide triphosphates (dNTPs) and removes 3' overhangs (Sambrook et al., 1989).

To prevent vector self-ligation, digested plasmid DNA was usually treated with shrimp alkaline phosphatase (Amersham Life Sciences) to remove the 5' phosphate group from the vector ends. In accordance with the manufacturers instructions, the digested vector was incubated with an excess of phosphatase (1 unit $\mu$g$^{-1}$) at 37°C for 1 hour. The vector was then incubated at 65°C for 15 min to inactivate the phosphatase and the DNA was extracted with phenol:chloroform:isoamyl alcohol 50:49:1 v/v, ethanol precipitated and resuspended in sterile distilled water, as described in section 2.3.1.1. After phosphatasing, the vector was self-ligated and a proportion of the ligation was transformed and plated, as described below. If there were lots of colonies on the plate, vector re-ligation was deemed to be a problem and the vector was re-phosphatased before further use.

2.3.5 DNA ligation and bacterial transformation
The vector and insert were usually incubated overnight at room temperature with T4 DNA ligase (Gibco BRL), according to the manufacturers instructions. Usually ligations with different vector:insert ratios were set up. All of the ligations had a final volume of 10$\mu$l. Ligations were transformed into E. coli XL1-Blue (Stratagene). Competent cells were prepared using the method of Miller (1994) and stored as 200$\mu$l aliquots in 10% glycerol (v/v) at −80°C until required. Electroporation was performed using a Bio-Rad gene pulser as described by Miller (1994), except the shocking chamber and cuvettes were cooled to −20°C before use and the cells were given a 1 hour period of outgrowth. In this procedure, 40$\mu$l of the competent cell suspension was mixed with 3$\mu$l of each ligation mixture, transferred to a 2mm cuvette (Bio-Rad) and positioned in the electroporator. A pulse of 12.5kV cm$^{-1}$ (2.5kV, 25$\mu$F, 200Ω) was applied across the cuvette and the cells were then
transferred to 1ml of SOC broth for a 1 hour period of outgrowth at 37°C. Aliquots of bacteria were also transformed with 100ng of pGem 7Zf+ (Promega) and 5μl of sterile distilled water as positive and negative controls respectively. These controls showed when the transformation system was not working or when contamination had occurred.

LB-amp plates were inoculated with 0.01μl of the positive control, 200μl of the negative control and 10μl and 100μl of the remaining transformation mixtures, and the plates were incubated at 37°C overnight. The remainder of the transformation mixtures were stored at 4°C. After incubation, the colonies on the positive control plate were counted to calculate the transformation efficiency (transformants per μg of DNA). Transformation efficiencies typically ranged from 1-6x10³. No colonies were usually seen on the negative control plate. Many colonies were seen on the plates where the ligation was successful. If individual colonies could be identified, 20 were picked and plasmid DNA was extracted from each (section 2.2.6). The plasmid DNAs were then analysed by restriction enzyme digestion and agarose gel electrophoresis to see if the plasmid of interest had been obtained (sections 2.3.1.2 and 2.3.2). If individual colonies could not be identified or if the plates did not contain enough colonies, an adjusted amount of the transformation mixture stored at 4°C was re-plated.

2.3.6 Making frozen stocks of bacterial cultures
Colonies of interest were inoculated into 5ml of LB-amp using a wire loop and grown overnight on a shaker at 37°C. Sterile 50% v/v glycerol was added to 2ml of the culture to a final volume of 10% v/v and an equal volume of the suspension was placed into 2 cryo-tubes. The cell suspensions were then frozen using a dry-ice/ethanol bath and stored at -80°C until required. For the genomic library, the transformation mixtures were transferred in 1ml aliquots to cryo-tubes. Glycerol was added to each of the aliquots and they were frozen and stored, as described above.

2.4 NUCLEIC ACID ANALYSIS
2.4.1 Southern blotting
Approximately 3μg of restriction enzyme digested genomic DNA (section 2.3.1.3) was electrophoresed overnight in 0.8% agarose gels with λ HindIII and φX174 HaeIII DNA size markers (section 2.3.2). The gels were then stained with ethidium bromide, photographed against a fluorescent ruler (section 2.3.2) and treated with depurinating (0.25M HCl), denaturing (1M NaCl and 0.5M NaOH) and neutralising (0.5M Tris and 3M NaCl, pH 7.5) solutions for 15-20 min each, as described by Sambrook et al. (1989). The DNA was transferred to Hybond-N membrane (Amersham) using the 10xSSC (1.5M sodium chloride and 150mM sodium citrate: Sambrook et al., 1989) alkaline transfer
method of Sambrook et al. (1989). The DNA was fixed to the membrane by exposure to
UV using an UV crosslinker (Amersham) with an energy setting of 700x100μJ cm⁻².

2.4.2 Northern blotting
Approximately 10μg of each total RNA sample and 3μg of RNA size markers (Promega)
were denatured at 50°C for 20 min in 1 volume of GFM (1.1M glyoxal, 78% v/v
formamide and 0.06xMOPS). A 0.13 volume of RNA loading buffer (4% v/v formamide,
5% v/v glycerol, 0.1xMOPS, 0.0825% w/v xylene cyanol and 0.0025% w/v bromophenol
blue) was added to each of the RNA samples immediately prior to loading. The samples
were then run on a gel containing 1.1% agarose v/v in 1xMOPS (20mM MOPS, 5mM
sodium acetate and 100μM EDTA pH 7) and electrophoresed in 1xMOPS. After running,
the lane containing the RNA size markers was excised from the gel, stained with
100ng ml⁻¹ ethidium bromide and photographed against a fluorescent ruler. The RNA
samples were then transferred to Hybond-N membrane (Amersham) using the 20xSSC
alkaline transfer method of Sambrook et al. (1989). The RNA was fixed to the membrane
by exposure to UV, as described above.

2.4.3 Colony blots
Bacterial stocks were plated onto LB-amp plates at a density of approximately 1000
colonies per plate and incubated overnight at 37°C. A circle of Hybond-N membrane
(Amersham) was placed onto the surface of each plate and left for 2 min. During this time
the filter and the agar were pierced in a distinct pattern with a sterile needle to allow filter
re-alignment later. The filter was then lifted from the plate and placed, colony side up, onto
filter paper to dry. The LB-amp plates were left at room temperature overnight to allow
colony re-growth and were subsequently stored at 4°C. When the filters were dry they
were placed onto a double thickness of filter paper soaked with denaturing solution (0.5N
NaOH and 1.5M NaCl) and left for 7 min. They were then transferred to filter paper
soaked in neutralising solution (1.5M NaCl and 0.5M Tris-HCl pH 7.4) and left for 3 min.
The neutralising step was repeated before the filters were washed in 2xSSC and placed
onto filter paper to dry. The DNA was fixed to the membrane by exposure to UV (section
2.4.1). The blots were then hybridised to a radio-labelled probe (section 2.4.4) and exposed
to x-ray film at -80°C. Positives on the x-ray film were traced back to individual colonies
on the plate at 4°C using the orientation marks. Colonies of interest were inoculated into
1ml LB-amp to enable a plasmid isolation procedure to be carried out (section 2.2.6) and
onto an LB-amp plate for storage. The plasmids were analysed by restriction enzyme
digestion (section 2.3.1.2) and gel electrophoresis (section 2.3.2).
2.4.4 Radio-labelling probes and hybridisation

Vector-free inserts were purified as described in section 2.3.3. Radio-labelled DNA probe fragments were generated from 100-300ng of insert using the random primer method of Feinberg and Vogelstein, (1983). The reaction contained 10-25mCi ml$^{-1}$ $\alpha$-$P^{32}$ dCTP and was incubated at room temperature for approximately 4 hours. The reaction was catalysed by the DNA polymerase I large (Klenow) fragment (Gibco BRL), according to the manufacturers instructions. After incubation, 90µg unsheared salmon sperm DNA was added to the reaction and the DNA was precipitated with ethanol to remove any unincorporated $P^{32}$-dCTP. The salmon sperm and probe DNA pellet was then resuspended in 500µl sterile distilled water, incubated at 100°C for 5 min and cooled on ice before being added to the hybridisation mixture.

The hybridisation solution contained 0.5M Na$_2$HPO$_4$ pH 7.5, 1mM EDTA and 7% w/v SDS (sodium dodecyl sulphate). Blots were prehybridised with 10-20ml of hybridisation solution for approximately 4 hours in a rotisserie dual hybridisation oven (Hybaid). The probe was then added to the solution and the blot was hybridised overnight. Usually blots were hybridised at high stringency (65°C), although occasionally blots were hybridised at low stringency (55°C). After hybridisation, the blot was subjected to four 10 min washes at the hybridisation temperature; it was washed twice with 1% w/v SDS and 0.2M Na$_2$HPO$_4$ pH 7.5 and twice with 1% w/v SDS and 0.1M Na$_2$HPO$_4$ pH 7.5 (based on Church and Gilbert, 1984). The blot was then removed from the hybridisation tube, wrapped in cling film and exposed at -80°C with an intensifying screen. Blots that were re-probed were stripped with boiling 1% w/v SDS as described by Sambrook et al. (1989), before being hybridised to the second probe.

2.4.5 DNA sequencing

Purified plasmid was isolated from bacteria as described in section 2.2.5 Automated sequencing reactions were carried out using an ABI PRISM dye terminator cycle sequencing reaction ready kit (Perkin-Elmer), according to the manufacturers instructions. Following amplification, the DNA was precipitated with 2 volumes of ethanol and a 0.1 volume of 3M sodium acetate, placed at 4°C for 4 hours and pelleted at maximum speed in a microfuge for 15 min. The supernatant was removed and the pellet was allowed to air dry. The dry pellet was supplied to the Protein and Nucleic Acids Laboratory (PNACL) at Leicester University where the DNA was analysed with an Applied Biosystems 337 or 337XL ABI automated sequencer. All sequences were analysed using the Wisconsin Package, Genetics Computer Group (Madison, USA) on the Leicester University mainframe computer. BLAST, FASTA, and TFASTA database searches were carried out.
Sequencing Primers were synthesised using an ABI 394 DNA synthesiser (Perkin-Elmer) at a concentration of 40\textmu M by PNACL (Leicester University). The primer was precipitated with one volume propan-2-ol, incubated at -20°C and pelleted in a microfuge at maximum speed for 10 min. The pellet was resuspended in 100\textmu l of sterile distilled water and the DNA concentration was determined using a spectrophotometer, as described in section 2.2.1. The primer concentration (mM) was then calculated using the molar extinction coefficients in the following equation:

\[
\text{Concentration (mM)} = \frac{\text{OD}_{260} \times \text{dilution}}{(A_n \times 15.4) + (T_n \times 8.8) + (G_n \times 11.7) + (C_n \times 7.7)}
\]

Where A_n, T_n, G_n, and C_n specify the number of respective bases in the primer sequence. Table 2.3 shows the sequence of all of the primers that were used during this work.

WORK CONDUCTED IN VILLEJUIF
The RT-PCR described in Chapter 3 and all of the work described in Chapter 5 was conducted at the Institut de recherches sur le Cancer, CNRS-UPR 1983, 7 rue guy Moquet-BP8, 94801, Villejuif, Cedex, France where I worked under the supervision of Dr. Gérard Pierron. The following sections describe the methods that I used in Villejuif which varied from those used in Leicester.

2.5 P. POLYCEPHALUM CULTURE AND NUCLEIC ACID ISOLATION
2.5.1 Culturing macroplasmodia
Genomic DNA was isolated from plasmodia at different stages of S-phase for the quantitative Southern blot analysis. Macroplasmodia ranging from 5-7cm in diameter were grown from microplasmodia that had been grown in axenic culture for two days, as described in Bénard and Pierron (1990). The microplasmodia were pelleted at 200\texttimes g for 2 min, the supernatant was removed and the cells were resuspended in 2 volumes of sterile distilled water. Several mls of the cell suspension was then deposited in the middle of a sheet Whatman filter paper located on top of a wire grid within a 9cm glass plate and the plate was incubated at 25°C for 90 min to allow the cells to fuse to form a single cell. After the 90 min incubation, the plasmodium was fed by placing several ml of SDM into the bottom of the plate and incubated at 25°C overnight.

After plasmodial fusion, the nuclei in the macroplasmodium are in a common cytoplasm and their mitotic cycles quickly synchronise. The second and third synchronous mitoses occur approximately 15 and 24 hours after the plasmodium is fed. After overnight incubation, a small section of the macroplasmodium was scraped from the plate at regular intervals, spread thinly onto a glass slide and examined under a phase-contrast microscope.
to see what stage of the mitotic cycle the nuclei were in. In the plasmodium, DNA synthesis begins directly after mitosis (Nygaard, 1960). Thus, the start of S-phase is defined as the observation of telophase. At the start of S-phase the plasmodium was incubated for a specific time, e.g. if a plasmodium that was 30 min into S-phase was required, the plasmodium was incubated for a further 30 min before being harvested for DNA isolation (section 2.5.3). As it was not always convenient to proceed with a DNA isolation procedure, a plasmodium was sometimes frozen in liquid nitrogen and stored at -20°C until required.

2.5.2 Treating a plasmodium with BromodeoxyUridine (BrdU)
Macroplasmodia were treated with BrdU as described in Bénard and Pierron (1990), using macroplasmodia that were grown as described above. At the start of S-phase, the filter paper containing the plasmodium was lifted from the grid in the glass plate and placed onto a sheet of filter paper that had been soaked with 100μg ml⁻¹ BrdU, 5μg ml⁻¹ fluorodeoxyuridine and 100μg ml⁻¹ uridine in a fresh plate. Fluorodeoxyuridine prevents thymidine incorporation by blocking production of the nucleotide, deoxythymidylate and uridine is added to counter the inhibition of RNA synthesis that is caused by giving a plasmodium fluorodeoxyuridine (Sachsenmaier and Rusch, 1964). The plasmodium was then incubated at 25°C for a specific time. During this time, the plasmodium takes up the BrdU, fluorodeoxyuridine and uridine and BrdU molecules are incorporated into the replicating DNA. The DNA that is replicated during the time the plasmodium is in the presence of BrdU is BrdU-substituted and has one ‘light’ parental strand and one ‘heavy’ newly-synthesised strand. In contrast, the DNA that is not replicated during this time has no BrdU substitution and has two ‘light’ parental strands. These two types of DNA are referred to as ‘heavy light’ (HL) and ‘light light’ (LL) DNA. After incubation, the plasmodium was harvested for DNA isolation or frozen until required, as above.

2.5.3 Isolation of genomic DNA from macroplasmodia
Genomic DNA was isolated from TU291 or M3CIV macroplasmodia, as described in Bénard and Pierron, (1990) although the composition of the homogenisation solution was modified to include 0.25M sucrose. A macroplasmodium was homogenised in homogenisation solution (0.25M sucrose, 0.1% v/v Nodinet p40, 10mM CaCl₂ and 10mM Tris-HCl) to release the nuclei. The homogenate was then filtered through a milk filter, to remove the poorly homogenised material, and the nuclei were washed in sterile distilled water. The nuclei were resuspended in lysis buffer (50mM Tris, 50mM NaCl and 25mM EDTA pH8) and treated with 100μg ml⁻¹ RNaseA and 250μg ml⁻¹ Proteinase K. The DNA was cleaned by successive extractions with phenol and chloroform:isoamyl alcohol (49:1 v/v), precipitated with 2 volumes of ethanol and a 0.1 volume of 3M sodium acetate pH
5.2 and pelleted. The supernatant was removed and the DNA pellet was dried slightly and resuspended in 1xTE (10mM Tris-HCl, 1mM EDTA pH 8.0). The DNA was quantified using a spectrophotometer (see section 2.2.1).

2.5.3.1 Embedding genomic DNA in agarose plugs
Genomic DNA in agarose plugs was used for 2D gel electrophoresis. Genomic DNA was embedded in agarose plugs using the method of Bénard et al. (1996). In this procedure, nuclei were isolated as described above, pelleted by centrifugation and resuspended in the homogenisation solution. This suspension was then mixed with 1% low-melting-point agarose solution at 42°C and 2mm diameter tubing was filled with the mixture and immersed in 1xTE at 4°C to solidify the agarose. The gel was later removed from the tubing and the agarose cylinders were treated with a solution containing 1mg ml\(^{-1}\) proteinase K, 1% v/v sarcosyl and 0.4M EDTA. The agarose cylinders were then washed repeatedly in 1xTE, treated with 100μg ml\(^{-1}\) RNaseA and washed again in 1xTE. The cylinders were stored in 1xTE at 4°C for several months. When needed, 8x2mm plugs were cut from the end of the agarose cylinder and the DNA was digested in the plugs and run on a 2D gel (sections 2.6.1.4 and 2.6.2). A plug usually contained 10μg of genomic DNA.

2.5.4 RNA isolation and RT-PCR
Two of the total RNA samples that were used for the RT-PCR were isolated in Leicester as described in section 2.2.2. These samples were the RNA from CL amoebae and the developing culture in which 56% of the cells were committed to development (Bailey et al., 1999; section 2.2.2; Table 2.1). Total RNA was isolated by G. Pierron from an LU352 macroplasmodium following solubilisation in guanidium hydrochloride and overnight centrifugation onto a CsCl cushion, as described in Pierron et al. (1989). The RNA pellet was resuspended in DEPC-treated water, ethanol precipitated and stored at \(-80°C\). RT-PCR was carried out using a kit from Stratagene, according to the manufacturers instructions; 10μg of total RNA was reverse transcribed from random hexamers with Moloney murine leukaemia virus reverse transcriptase. The resulting RT-cDNA samples had a volume of 50μl and were stored at \(-20°C\) and used repeatedly. A 1μl volume of each RT-cDNA sample was used to template each PCR reaction using primers specific for the gene of interest (Table 2.3). The required primers were synthesised by Genset oligos and diluted to a 10mM concentration. The PCR reactions contained 1μl of each primer and were catalysed by Taq+ long polymerase (Stratagene), according to the manufacturers instructions. For the PCR reactions, the reaction mixture was heated to 72°C for 5 min. The Taq polymerase was then added to the reaction mixture which was covered with mineral oil and subjected to 30 amplification cycles: 91°C for 30 sec, 56°C for 30 sec and
72°C for 45 sec. After amplification, one fifth of each PCR reaction (10µl) was electrophoresed on a 2% agarose gel (section 2.6.2).

2.6 RESTRICTION ENZYME DIGESTION AND AGAROSE GEL ELECTROPHORESIS

2.6.1 Restriction enzyme digestion

Restriction digests were carried out using Promega restriction enzymes, in accordance with the manufacturers instructions. Most of the restriction enzymes were at a concentration of 40-80 units µl⁻¹, which allowed large quantities of enzyme units to be added to a digest without glycerol inhibiting the reaction.

2.6.1.1 Plasmid digests

Plasmid DNA, isolated at Leicester University (section 2.2.5) was digested to obtain vector-free purified inserts that were used as probes. Approximately 2µg of plasmid DNA was digested with an excess of enzyme (10 units µg⁻¹) at 37°C for 2hr in a final volume of 30µl. The digests were then loaded straight onto an agarose gel (section 2.6.2).

2.6.1.2 Genomic DNA digests

Genomic DNA from macroplasmodia (section 2.5.3) was digested for Southern blotting. Approximately 3µg of genomic DNA was digested with an excess of enzyme (10 units µg⁻¹) at 37°C for approximately 4 hours in a final volume of 60µl. The digests were then loaded straight onto the gel (section 2.7.1).

2.6.1.3 Digesting genomic DNA from BrdU-treated macroplasmodia

It was usual to digest all of the genomic DNA isolated from a BrdU-treated macroplasmodium at one time (approximately 100-200µg; section 2.5.2 and 2.5.3). The DNA was digested with an excess of enzyme (10 units µg⁻¹) at 37°C for approximately 4.5 hours in a final volume of 2ml. Additional enzyme (10 units µg⁻¹) was added to the digest 1.5 hours and 3 hours from the start of incubation. The DNA was then run on a CsCl gradient to separate the LL and the HL DNA (section 2.7.2).

2.6.1.4 Digesting genomic DNA in agarose plugs

Genomic DNA in agarose plugs (section 2.5.3.1) was digested for 2D gel electrophoresis as described in Bénard et al. (1996). An 8x2mm agarose plug, containing approximately 10µg of genomic DNA was placed into an eppendorf containing 1ml of the appropriate restriction enzyme buffer and incubated for 1hr at room temperature. The agarose plug was then transferred to a fresh eppendorf and digested with an excess of enzyme (100-200 units) at 37°C for 6 hours in a final volume of 200µl. Additional enzyme (100-200 units)
was added to the digest 2 hours and 4 hours from the start of the incubation. The agarose plug was then removed from the eppendorf, washed once in 1xTE and run on a 2D gel (section 2.7.3). When double enzyme restriction digests were required with enzymes that do not work efficiently in the same buffer, the DNA was digested as described above and left to soak in 1xTE overnight before the second restriction enzyme digest was performed.

2.6.2 Agarose gel electrophoresis

Gels containing 0.8-1.5% agarose v/v in 1xTAE (40mM Tris acetate and 2mM EDTA: Sambrook et al., 1989) were prepared and electrophoresed in 1xTAE. A 0.1 volume of loading buffer (0.25% w/v bromophenol blue in 40% w/v sucrose: Sambrook et al. 1989) was added to each of the DNA samples immediately prior to loading. Generally λ HindIII and/or φX174 HaeIII DNA size markers were included on the gels. The running buffer contained ethidium bromide at a concentration of 300ng ml⁻¹.

For 2D gel electrophoresis, gels containing 0.4% and 1.1% agarose v/v in 1xTBE (90mM Tris-Cl, 90mM boric acid and 2mM EDTA pH 8; Sambrook et al., 1989) were prepared and electrophoresed in 1xTBE. λ HindIII DNA size markers were included on the 0.4% gels.

2.6.3 Isolating restriction fragments of interest

Plasmid DNA isolated in Leicester (section 2.2.5) was digested as described in section 2.6.1.1. The digest was divided between several lanes on a 0.8% agarose and run with λ HindIII and φX174 HaeIII DNA size markers (section 2.6.2). After running, the DNA was visualised briefly under UV illumination and the restriction fragment of interest was located. The slice containing the restriction fragment was excised from the gel and the DNA was separated from the agarose by filtration through glass wool. In this procedure, the bottom of an eppendorf was pierced with a needle and a small plug of glass wool was placed into the bottom of the tube over the hole. The agarose slice was then placed into the eppendorf, which was put into a larger tube and centrifuged at approximately 800xg for 5 min. The pierced eppendorf was then discarded and the elute was transferred from the larger tube to a fresh eppendorf tube. The DNA was precipitated with 2 volumes of ethanol and a 1 in 10 volume of sodium acetate pH 5.2. The DNA was incubated at -20°C for 45 min and pelleted at maximum speed in a microfuge for 15 min. The supernatant was then removed and the pellet was washed in 80% v/v ethanol. After re-pelleting the supernatant was removed and the pellet was dried before being resuspended in 1xTE. A small amount of the purified insert was run on a 0.8% agarose gel with φX174 HaeIII DNA size markers of known concentration. The gel was photographed under UV illumination and the DNA
concentration was estimated by comparing the intensity of the insert band to the intensity of several DNA size marker bands.

2.7 NUCLEIC ACID ANALYSIS

2.7.1 Southern blotting

3μg of genomic DNA was digested (section 2.6.1.2) and run on a 0.8% agarose gel with λ HindIII and φX174 HaeIII DNA size markers (section 2.6.2). After running, the gel was photographed under UV illumination against a fluorescent ruler. The gel was then treated with depurinating (250mM HCl), denaturing (600mM NaCl and 400mM NaOH) and neutralising (500mM Tris and 1.5M NaCl pH 7.6) solutions for 15 min, 30 min and 30 min respectively. The DNA was transferred to Gene Screen plus membrane (NEN) using VacuGene transfer apparatus (Pharmacia) with 10xSSC (1.5M sodium chloride and 150mM sodium citrate: Sambrook et al., 1989). Immediately after transfer, the membrane was subjected to two 1 min washes in 400mM NaCl and 200mM Tris in 1xSSC and blotted dry. The DNA was fixed to the membrane by baking at 80°C in a vacuum oven for 1 hour.

2.7.2 Separating the BrdU-substituted non-BrdU substituted DNA

The normal, LL DNA was separated from the density shifted, HL DNA (section 2.5.2) on a CsCl gradient using a protocol devised by G. Pierron (unpublished) which was based on the protocol of Pierron et al. (1984). The genomic DNA (100-200μg) from a BrdU-treated macroplasmodium was digested with a restriction enzyme (section 2.6.1.3). After digestion, the DNA was diluted to 4ml in a solution containing 1xTE, 0.25% v/v sarcosyl and 300ng ml⁻¹ ethidium bromide. 4.95g of solid CsCl was dissolved into the solution, which was then centrifuged for 16 hours at 70,000 rpm in a vertical rotor (NVT 90) in a Beckman LE-80 ultracentrifuge. After centrifugation, the LL and HL DNA bands were visualised using a hand held UV lamp. The LL and the HL DNA fractions were extracted separately from the centrifuge tube using a hypodermic syringe with a 21-gauge needle and transferred to fresh tubes. The two fractions were then diluted with 2 volumes of 1xTE and aliquoted into 500μl samples. The DNA was precipitated in 2 volumes of ethanol, incubated at −20°C for 1 hour and pelleted at maximum speed in a microfuge for 15 min. The supernatant was removed and the pellet was resuspended in 1xTE.

To estimate the DNA concentration, a small amount of each DNA fraction was run on a 0.8% agarose gel with λ HindIII and φX174 HaeIII DNA size markers (section 2.6.2). Then, approximately 5μg of the digested LL and HL DNA from a BrdU-treated plasmodium was run on a 0.8% agarose gel (section 2.6.2). The DNA transferred to membrane, as described above.
2.7.3 2D gel electrophoresis

2D gel electrophoresis was carried out as described in Brewer and Fangman (1987). Approximately 10μg of genomic DNA imbedded in an agarose plug (section 2.5.3.1) was digested with restriction enzyme(s) (section 2.6.1.4). The agarose plug was loaded onto a 0.4% agarose gel and run overnight (section 2.6.2). The following day the gel was stained with 300ng ml$^{-1}$ ethidium bromide and the lane containing the λ HindIII DNA size markers was photographed under UV illumination. The lane containing the digested genomic DNA was then excised from the gel and placed horizontally across the top of a clean gel bed in the position the lanes usually occupy. Melted 1.1% agarose containing 300ng ml$^{-1}$ ethidium bromide was then poured into the gel bed and allowed to solidify around the 0.4% gel slice to make the second gel (section 2.6.2). This gel was run at 4°C at high voltage (3V cm$^{-1}$) for 15 hours in buffer containing ethidium bromide at a concentration of 300ng ml$^{-1}$. The DNA was transferred to membrane, as described in section 2.7.1. The position of the linear DNA was marked on the dry blot using a hand held UV lamp.

2.7.4 Radio-labelling probes and hybridisation

Radio-labelled probe fragments were generated from 100-300ng of vector-free purified insert in a random-primed labelling reaction using a NEN kit. The reaction contained 25-50 mCi ml$^{-1}$ of α-32P dCTP and was incubated at room temperature for at least 1 hour. After incubation, approximately 100μg of unsheared salmon sperm DNA was added to the reaction, which was incubated at 100°C for 10 min, cooled on ice for 10 min and added to the hybridisation mixture. Blots were pre-hybridised with 10ml of hybridisation solution (1% w/v BSA [bovine serum albumin], 7% w/v SDS, 1mM EDTA and 0.5M phosphate buffer pH 7.4 [1M NaH$_2$PO$_4$ and 1M Na$_2$HPO$_4$]: Church and Gilbert, 1984) at 65°C in a rotisserie dual hybridisation oven (Hybaid) for 1-4 hours. The probe was then added to the mixture and the blot was hybridised for a further 15 hours. After hybridisation, the blot was subjected to four 10 min washes at the hybridisation temperature with 50ml of wash solution (40mM phosphate buffer, 1mM EDTA and 1% w/v SDS: Church and Gilbert, 1984). The blot was then removed from the hybridisation tube, wrapped in cling film and exposed at -80°C with 2 intensifying screens. Gene dosage blots were exposed to a special screen, which was scanned using a STORM phosphorimager (Molecular dynamics). The band intensities were measured using Image Quant version 1.2.
Table 2.1: The distribution of cell types in the cultures of apogamic developing cells from which total RNA was isolated (Bailey et al., 1999)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amoebae (%)</th>
<th>Uninucleate developing cells (%)</th>
<th>Multinucleate developing cells (%)</th>
<th>Total developing cells (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LU352</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1%</td>
<td>98.9</td>
<td>1.1</td>
<td>0</td>
<td>1.1</td>
</tr>
<tr>
<td>4%</td>
<td>95.9</td>
<td>3.4</td>
<td>0.7</td>
<td>4.1</td>
</tr>
<tr>
<td>10%</td>
<td>90.4</td>
<td>9.3</td>
<td>0.3</td>
<td>9.6</td>
</tr>
<tr>
<td>11%</td>
<td>89.3</td>
<td>10.4</td>
<td>0.3</td>
<td>10.7</td>
</tr>
<tr>
<td>22%</td>
<td>77.9</td>
<td>21.2</td>
<td>0.9</td>
<td>22.1</td>
</tr>
<tr>
<td>23%</td>
<td>77.3</td>
<td>18.8</td>
<td>3.9</td>
<td>22.7</td>
</tr>
<tr>
<td>38%</td>
<td>62.3</td>
<td>32.5</td>
<td>5.2</td>
<td>37.7</td>
</tr>
<tr>
<td>39%</td>
<td>61.0</td>
<td>34.7</td>
<td>4.3</td>
<td>39.0</td>
</tr>
<tr>
<td>56%</td>
<td>44.0</td>
<td>46.1</td>
<td>9.9</td>
<td>56.0</td>
</tr>
<tr>
<td>Mi</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ma</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

CL and LU352: amoebae; 1-56%: apogamic developing cultures of CL; Mi: CL microplasmodia, Ma: CL macroplasmodia. See section 2.2.2 for details of how the distribution of cells types was determined.
Table 2.2: The distribution of cell types in the cultures of heterothallic developing cells from which total RNA was isolated (Barber, 1998)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amoebae (%)</th>
<th>Fusion cells (%)</th>
<th>Zygotes (%)</th>
<th>Multi-nucleate developing cells (%)</th>
<th>Total developing cells (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH508</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LU352</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2%</td>
<td>97.7</td>
<td>0.8</td>
<td>1.5</td>
<td>0</td>
<td>2.3</td>
</tr>
<tr>
<td>4%</td>
<td>95.7</td>
<td>2.1</td>
<td>2.2</td>
<td>0</td>
<td>4.3</td>
</tr>
<tr>
<td>17%</td>
<td>82.8</td>
<td>3.7</td>
<td>13.5</td>
<td>0</td>
<td>17.2</td>
</tr>
<tr>
<td>24%</td>
<td>76</td>
<td>1.2</td>
<td>14.4</td>
<td>8.4</td>
<td>24</td>
</tr>
<tr>
<td>38%</td>
<td>62</td>
<td>4.1</td>
<td>25.7</td>
<td>8.2</td>
<td>38</td>
</tr>
<tr>
<td>41%</td>
<td>59</td>
<td>4.9</td>
<td>29.1</td>
<td>7</td>
<td>41</td>
</tr>
<tr>
<td>Mi</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ma</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

CH508 and LU648: amoebae; 2%–24%: developing cultures obtained by crossing CH508 and LU648 amoebae; 38% and 41%: developing cultures obtained from ten-strain matings (see below); Mi: microplasmodia obtained by crossing CH508 and LU648 amoebae; Ma: macroplasmodia obtained by crossing CH508 and LU648 amoebae. See section 2.2.1 for details of how the distribution of cells types was determined.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse</td>
<td>5’- AAC AGC TAT GAC CAT G -3’</td>
</tr>
<tr>
<td>-40</td>
<td>5’- GTT TTC CCA GTC ACG AC -3’</td>
</tr>
<tr>
<td>proAd1</td>
<td>5’- GGC CAC GTT ATC GGA GCA GC -3’</td>
</tr>
<tr>
<td>proAm1</td>
<td>5’- GGA CAA GGA CCT CCC ACT GG -3’</td>
</tr>
<tr>
<td>proPdl</td>
<td>5’- ACC CCC GCC AAT GTC TTT GC -3’</td>
</tr>
<tr>
<td>proPml</td>
<td>5’- CTC AAT GAG GTA GTC GGC TAG -3’</td>
</tr>
<tr>
<td>redEd2</td>
<td>5’- AAT GAC ATC CTG GAG GCC AG -3’</td>
</tr>
<tr>
<td>redEm2</td>
<td>5’- CTA GGA CAC GGT ACT TGT GG -3’</td>
</tr>
<tr>
<td>redE1</td>
<td>5’- GTG GAA AAC ACG CTG TTG -3’</td>
</tr>
<tr>
<td>redE2</td>
<td>5’- TAT GTG CGG AAG AGA GTG G -3’</td>
</tr>
<tr>
<td>A181</td>
<td>5’- TCT CGA ATC CAC CAA GAC -3’</td>
</tr>
<tr>
<td>A182</td>
<td>5’- TCT TGT AGT CTG CCT CCA TC -3’</td>
</tr>
</tbody>
</table>
CHAPTER 3: THE IDENTIFICATION AND ANALYSIS OF RedE

3.1 INTRODUCTION
Studies of P. polycephalum strains which carried npf mutations led Bailey et al. (1992a) to propose that plasmodium development is controlled by a specific set of genes that are expressed during the APT (section 1.3.2). In order to find these genes, Bailey et al. (1992a) constructed a cDNA library with RNA isolated from a culture which contained a high proportion of developing cells, and enriched the library for genes that were expressed during development by subtraction (section 1.3.3.1). The subtracted cDNA library, ML8S, was then screened for genes that were differentially expressed. The screening identified 150 colonies which contained cDNAs representing differentially expressed genes (section 1.3.3.2). The expression pattern of 12 of these cDNAs was then examined by northern blotting (section 1.3.3.2). The northern blotting analysis led to the identification of a novel class of genes that are expressed primarily during development; they are not expressed in amoebae, are expressed at relatively high levels in developing cells and are expressed at low levels in plasmodia (section 1.3.3.3). These genes were called red (regulated in development) genes and the two that were identified were named redA and redB (section 1.3.3.3; Bailey et al., 1999). The first aim of this work was to identify more red genes. The second aim was to characterise any isolated gene(s).

RESULTS
3.2 THE IDENTIFICATION AND PRELIMINARY ANALYSIS OF POTENTIAL red GENES
My first aim was to identify some more red genes. I continued to analyse the 150 frozen stocks that were made during the first round library screen of the subtracted cDNA library, ML8S, as only 12 of these were analysed in detail (section 1.3.3.2). Instead of screening the frozen stocks with subtracted probes, as in Bailey et al. (1999; section 1.3.3.2), a different approach was taken, that is detailed below.

Fifty of the 150 frozen stocks were chosen at random. In order to view the inserts in these frozen stocks, plasmid DNA was extracted from each and digested with XhoI and SstI, to excise the inserts from the vectors (sections 2.2.5 and 2.3.1.2). The digests were then run on a 1% agarose gel (data not shown) and the DNA was stained with ethidium bromide and examined under UV illumination (section 2.3.2). Some of the frozen stocks did not appear to contain any plasmids with inserts, while others contained one or more inserts. The occurrence of multiple inserts was not surprising, because earlier analysis had shown that many of the 150 colonies contained bacteria from more than one colony (section 1.3.3.2; Bailey et al., 1999). It was also possible that the enzymes had cut within an individual insert to produce two or more bands from a single insert. I chose 16 of the
frozen stocks which contained at least one insert to analyse further. In order to confirm that the 16 frozen stocks contained inserts, plasmid DNA was again digested and run on a 1% gel (data not shown), as described above. The results showed that the 16 frozen stocks contained a total of 40 inserts. The insert sizes were measured and each was given a unique name. For example, the frozen stock, D4 contained two inserts, one of 600bp and one of 790bp; these inserts were named D4/600 and D4/790 respectively.

It seemed likely that some of the 40 inserts would represent cDNAs that had already been isolated from ML8S (e.g. redA and redB). In order to avoid analysing cDNAs that had already been characterised, the eleven ‘known cDNAs’ were cross-hybridised to the 40 inserts. Inserts that hybridised to any of the known cDNAs were not investigated further. Plasmid DNA from the 16 frozen stocks was digested and electrophoresed, as above, and the gel was blotted (section 2.4.1). To reduce the number of separate screens required, the blot was probed with two of the known cDNAs in a single hybridisation (section 2.4.4). Eleven of the 40 inserts were eliminated from further analysis because they cross hybridised to one or more of the known cDNA clones (data not shown). This left 29 inserts of interest contained within 12 of the 16 frozen stocks.

Several of the 12 frozen stocks were chosen for further analysis. Bacteria from each stock were streaked onto an LB-amp plate to give individual colonies. From each plate, several colonies were picked at random and plasmid DNA was extracted from each, digested and run on a 1% agarose gel (data not shown; sections 2.2.6, 2.3.1.2 and 2.3.2). Five individual colonies that contained plasmid with a single insert were identified. The clones were named D2/650, D2/750, D11/1100, D15/1060, D15/1150 and frozen stocks were made of each (section 2.3.6). A sixth colony that appeared to contain plasmid with two inserts, A18/375 and A18/630 was also frozen. In order to see if the A18 stock contained a single insert that had been cut into two by XhoI or SstI, several test digests were set up and the digests were run on a 1% agarose gel (data not shown; sections 2.3.1.2 and 2.3.2). The results showed that the plasmid contained a single insert; a KpnI and SstI digest released an insert of approximately 1020bp. This clone was renamed A18/1020.

To determine whether any of the six clones represented red genes, their expression was analysed by northern blotting (section 2.4.2). Purified cDNA inserts were used to probe northern blots containing 10µg of total RNA isolated from CL amoebae, microplasmodia and a culture which contained a high proportion of developing cells (56%; Table 2.1; Bailey et al., 1999). The northern blots were probed at high stringency (data not shown). Three inserts, D2/650, D2/750 and D15/1060 did not produce any signal on the northern blots, presumably for the reasons discussed in section 1.3.3.2, and D15/1150 represented a constitutively expressed gene; these inserts were not analysed further. The A18/1020 clone
hybridised strongly to the RNA from developing cells, and very weakly to the RNA from microplasmodia, while D11/1100 hybridised only to the RNA from developing cells. A18/1020 hybridised to a mRNA of approximately 6500nt and D11/1100 hybridised to a mRNA of approximately 1200nt, indicating the D11/1100 cDNA is almost full length. These results suggested that D11/1100 and A18/1020 may represent red genes; the rest of this chapter describes the analysis of D11/1100 and Chapter 4 describes the analysis of A18/1020. The search for more red genes was discontinued while these cDNAs were characterised.

As part of the preliminary analysis of D11/1100, a Southern blot containing *EcoRI*, *BgII*, *HindIII*, and *EcoRV* digested CL genomic DNA was probed with the cDNA (data not shown; sections 2.3.1.3, 2.4.1 and 2.4.4). After hybridisation, a single (major) band could be seen in each lane which suggested that D11/1100 represents a single-copy gene.

### 3.3 SEQUENCING THE D11/1100 cDNA

An essential part of the analysis of D11/1100 was to determine the sequence of the cDNA. The cDNA was sequenced with the universal ‘reverse’ and ‘−40’ sequencing primers (Table 2.3), which have binding sites either side of the multiple cloning site within *pBluescript* II KS (Stratagene). The cDNAs were cloned such that the reverse primer amplified the sense strand from a 5’ direction, and the −40 primer amplified the antisense strand from a 3’ direction (section 1.3.3.1; Fig. 3.1a). Automated sequencing reactions were performed as described in section 2.4.5. Each primer gave data for approximately 300bp of sequence; a total of 600bp of sequence had been determined (Fig. 3.1a). As the D11/1100 cDNA is 1100bp in length, approximately 500bp of sequence was still missing.

#### 3.3.1 Subcloning the D11/1100 cDNA

In order to obtain the complete sequence of D11/1100, the cDNA was subcloned and some of the subclones were sequenced. Purified plasmid was digested with *XhoI* and *SstI* to release the cDNA insert, which was gel purified (sections 2.3.1.1 and 2.3.3). A restriction enzyme that cut the D11/1100 insert into suitably sized restriction fragments then had to be identified. Ideally the restriction fragments would be small enough to be fully sequenced with the reverse and the −40 sequencing primers (600bp or less), but large enough to carry a reasonable amount of sequence (100bp or more). Purified insert was digested with several restriction enzymes that had a 4bp recognition site (section 2.3.1.2). Because these enzymes cut DNA more frequently than 6bp cutters, it was thought that they would have more chance of cutting within the 1100bp insert. The digests were then electrophoresed on a 1% agarose gel (data not shown; section 2.3.2). The electrophoresis identified *Sau3AI* as a suitable enzyme, because it cut the DNA into several pieces. A total of seven bands could be seen on the gel, ranging from approximately 50-600bp in size. When all the bands were
added together they gave a size that was greater than the insert size of 1100bp, which indicated that some of the bands represented partially digested restriction fragments.

Purified insert was digested with Sau3AI and the ends of the restriction fragments were blunted with the Klenow fragment of E. coli DNA polymerase (section 2.3.4). Meanwhile, pBluescript II SK (Stratagene) was linearised with EcoRV, gel purified and treated with alkaline phosphatase (sections 2.3.1.1, 2.3.3 and 2.3.4). The Sau3AI restriction fragments were ligated into the pBluescript, and one third of the ligation mixture was transformed into E. coli XL1-blue and plated onto LB-amp plates (section 2.3.5). After overnight incubation many colonies could be seen on the plates. Twenty colonies were picked at random and plasmid DNA was extracted from each (section 2.2.6). The plasmids were digested with XhoI and SstI to release the inserts and the digests were electrophoresed on a 1% agarose gel (data not shown; sections 2.3.1.2 and 2.3.2). The electrophoresis showed that the cloning had been successful; every plasmid contained an insert, and there were nine differently sized inserts that ranged from approximately 40-600bp. Frozen stocks were made of colonies that contained the nine clones (section 2.3.6), and they were named subclones 1-9. Subclone 1 contained the largest insert and subclone 9 contained the smallest. When the D11/1100 cDNA was first digested with Sau3AI and electrophoresed, only seven bands were seen. Yet nine Sau3AI restriction fragments were cloned in the subcloning experiments. It seems likely that the two extra cloned fragments represent partially digested restriction fragments, that could not be seen on the first gel because they were either absent or low in concentration.

3.3.2 Sequencing D11/1100 ‘subclone 1’

The first subclone to be sequenced was subclone 1, because it contained the largest insert. The sequencing was carried out with the universal reverse and −40 sequencing primers as described in section 3.3 (section 2.4.5). The reverse primer gave data for approximately 550bp of sequence, and the −40 primer gave data for approximately 300bp of sequence (Fig. 3.1b). When the reverse and the −40 sequences were compared, there was a 200bp region of overlap (Fig. 3.1b) which showed that the complete sequence of subclone 1 had been determined. The clone was 636bp in length.

In order to see where subclone 1 was located within the cDNA, its sequence was compared to the reverse and the −40 sequence of the D11/1100 clone. The comparisons showed that the subclone 1 sequence overlapped with both the reverse and the −40 cDNA sequence, indicating the complete sequence of the D11/1100 cDNA had been obtained. The cDNA clone is 1102bp in length and has 18bp of poly(A)+ tail at its 3’ end (Fig. 3.2; EMBL accession number AJ297387). Two of the possible reading frames have stop codons scattered throughout. The other reading frame starts with the third base of the cDNA and
has the stop codon, (TAG), located 930bp into the sequence (Fig. 3.2); this must be the open reading frame (ORF). Thus, the cDNA clone consists of 929bp of coding sequence, 152bp of 3' untranslated region and 18bp of poly(A)+ tail. The cDNA is not full length because the start codon, ATG, is missing (Fig. 3.2). The deduced protein sequence contains 309 amino acids (aa's; Fig. 3.2). The analysis of the D11/1100 DNA and deduced protein sequence is continued in sections 3.7 and 3.8.

3.4 THE EXPRESSION OF D11/1100 DURING THE AMOEBAL-PLASMODIAL TRANSITION

Northern blotting suggested that D11/1100 could represent a red gene (section 3.2). In order to see if this were the case, a more detailed analysis of its expression was carried out using northern blotting and reverse transcription-PCR (RT-PCR).

3.4.1 Northern analysis

The northern analysis differed from that described in section 3.2 because more RNA samples were used. The northern blot contained total RNA from (i) CL amoebae (ii) CL microplasmodia (iii) CL macroplasmodia and (iv) a variety of CL developing cultures, each of which contained a different proportion of cells that were committed to apogamic development, ranging from 1-56% (1-56%; Table 2.1; Bailey et al., 1999; section 2.2.2). The developing culture RNA samples that were used to characterise the expression of redA and redB (section 1.3.3.3) were in this study, so that the expression of D11/1100 could be directly compared to that of the other red genes. Very poor signal was detected when northern blots containing 10µg of total RNA were probed for D11/1100, yet signal was detected when the blots contained 20µg of total RNA. This suggests the D11/1100 mRNA is low in abundance. Thus, a northern blot was made (section 2.4.2) which contained 20µg of each RNA sample. The blot was probed at high stringency (65°C; section 2.4.4) with the cDNA clones for D11/1100, proP (LAV1-5: Binette et al., 1990) and actin (pPpA35: Hamelin et al., 1988; Fig. 3.3). The probing was repeated several times using different northern blots, and the results shown in Figure 3.3 are representative of those obtained.

D11/1100 hybridised to two differently-sized mRNAs of 1300nt and 1150nt in length (Fig. 3.3a). Thus, the 1102bp cDNA is missing either 200bp or 50bp. The D11/1100 message could not be detected in the RNA from amoebae or the developing culture which contained 1% committed cells (Fig. 3.3a), but was detected in the RNA from all of the other developing cell cultures. The transcript levels increased with the percentage of committed cells to reach a maximum in the 39% and 56% samples (Fig. 3.3a). The D11/1100 transcript was detected at lower levels in the RNA from microplasmodia and macroplasmodia (Fig. 3.3a). The expression of D11/1100 is very similar to that of redA and redB; thus, the cDNA appears to represent a red gene.
Some control experiments were carried out before D11/1100 was finally designated a red gene; the blot was probed for actin and proP to allow the expression pattern of D11/1100 to be directly compared to that of a constitutively expressed and plasmodium-specific gene. As expected, actin and proP hybridised to 1400nt and 520nt mRNAs respectively (Pallotta et al., 1986; Binette et al., 1990). All of the RNA samples contain easily detectable actin transcript which is not significantly degraded (Fig. 3.3c). The plasmodial RNA samples contained less actin mRNA, and it is possible that these samples were underloaded or transferred badly. As expected, the proP mRNA was absent in the RNA from amoebae and the developing cell culture in which 1% of the cells were committed to development (Fig. 3.3b; Bailey et al., 1999). The proP message was detected in the RNA from all of the other developing cultures, and the amount of proP mRNA increased with the percentage of committed cells (Fig. 3.3b). The plasmodial RNAs contain the most proP mRNA (Fig. 3.3b). The actin control shows that the blot contains intact amoebal RNA. As the expression of D11/1100 cannot be detected in this sample, the cDNA cannot represent a constitutively expressed or amoeba-specific gene. The proP control shows that D11/1100 does not represent a plasmodium-specific gene, as the expression of D11/1100 is clearly different to that of proP, despite the possible underloading of the plasmodial RNA samples. These controls showed that D11/1100 does not represent a constitutively expressed, amoebal-specific or plasmodium-specific gene. Since the expression pattern of D11/1100 was the same as that of redA and redB, D11/1100 was designated a red gene and renamed redE.

Obtaining the northern data was problematic because the redE mRNA was of low abundance. In order to confirm that redE is a developmentally regulated gene that is not expressed in amoebae, its expression was analysed using RT-PCR. This technique was chosen because it can detect low abundance messages.

3.4.2 RT-PCR
In the RT-PCR experiments, 10μg of each total RNA was reverse transcribed to make an RT-cDNA sample, which was stored at −80°C and used repeatedly (section 2.5.4). The RT-cDNA samples were used to template a PCR reaction using primers specific for the gene of interest (section 2.5.4). The PCR reactions were then run on an agarose gel so that the amplification products could be visualised and sized (section 2.6.2). Specific amplification products are only generated if the gene of interest is expressed in the cell type from which the RNA was extracted. If the gene of interest is not expressed, there will be no amplification product as the primers will not have any cDNA molecules to amplify from. This type of RT-PCR does not give any indication of how abundant a specific mRNA is, it just indicates that a specific mRNA is present in the total RNA of a particular
cell type. The aim of the analysis was to confirm that redE was a developmentally regulated gene.

Total RNA (10μg) from (i) CL amoebae (ii) the developing cell culture containing 56% committed cells (56%: Table 2.1; Bailey et al., 1999) and (iii) LU352 microplasmodia was reverse transcribed (section 2.5.4). Two primers were designed that had binding sites within the redE cDNA, redEd2 and redEm2 (positions 215-234 and 913-932 respectively: Fig. 3.2; Table 2.3). It was predicted that redEd2 and redEm2 would amplify a 718bp region of the cDNA in a PCR reaction (Fig. 3.2). P. polycephalum genes have small introns interspersed with exons at regular intervals, e.g. proA, proP, redA and redB (Binette et al., 1990; Bailey et al., 1999). Thus, it seemed likely that redE would contain intron(s) between the redEd2 and redEm2 binding sites. It was thought that these intron(s) would be detected when genomic DNA was amplified with redEd2 and redEm2 by PCR, because the amplification product would contain the intron(s) and be larger than 718bp. However, when such a PCR reaction was set up, the amplification product was 718bp in length (data not shown). This shows that there are no introns within the genomic region of redE between the redEd2 and the redEm2 binding sites.

Some control experiments were designed to check that the PCR products were amplified from RT-cDNA molecules and not genomic DNA, because the contamination of total RNA by genomic DNA can sometimes be a problem. If redE had contained intron(s) within the 718bp region, a direct way telling whether a PCR product had been amplified from RT-cDNA or genomic DNA would have been available. When the PCR reactions were analysed by electrophoresis, the product amplified from RT-cDNA would not contain any introns and would generate a 718bp ‘intron-’ band. While product amplified from genomic DNA would generate a larger, ‘intron+’ band. As there were no introns located between the redEd2 and the redEm2 binding sites within redE, the amoebal-specific and the plasmodium-specific profilin genes, proA and proP were used as controls. These genes were suitable because they contained introns (Binette et al., 1990) which provided a means of telling whether contaminating genomic DNA was present in the samples. In addition, they had a known pattern of expression (Binette et al., 1990; Bailey et al., 1999) which allowed the reliability of the RT-PCR to be checked; proA should only be detected in the RT-cDNA from amoebae and the developing cell culture and proP should only be detected in the RT-cDNA from the developing cell culture and plasmodia (Bailey et al., 1999). Primer pairs were designed for proA, proAd1 and proAm1, and proP, proPd1 and proPm1 (Table 2.3). Both of the primer pairs had binding sites that were either side of the second introns within proA and proP; proAd1 and proAm1 spanned a 58bp intron, and proPd1 and proPm1 spanned a 69bp intron (see Binette et al., 1990). Thus, the proA primers amplified
a 465bp intron+ PCR product and a 407bp intron- product. While the proP primers amplified a 284bp intron+ PCR product and a 215bp intron- product.

Eight PCR reactions were set up (section 2.5.4): (1) The RT-cDNA from amoebae, (2) the RT-cDNA from the developing culture and (3) the RT-cDNA from plasmodia, was amplified with the redE primers. (4) Genomic DNA was amplified with the proA primers. (5) The RT-cDNA from amoebae, (6) the RT-cDNA from the developing culture and (7) the RT-cDNA from plasmodia was co-amplified with the proA and the proP primers. (8) Genomic DNA was amplified with the proP primers. After amplification, the PCR reactions were run on a 2% agarose gel (Fig. 3.4; section 2.6.2).

When proA was amplified from genomic DNA, it produced an intron+ band of 465bp (lane 4: Fig. 3.4). A 407bp proA PCR product was amplified from the RT-cDNA of amoebae and the developing cell culture, but not plasmodia (lanes 5, 6 and 7 respectively: Fig. 3.4). Thus, proA is expressed in amoebae and the developing culture, but not in plasmodia. When proP was amplified from genomic DNA, it produced and intron+ band of 284bp (lane 8: Fig. 3.4). A 215bp proP PCR product was amplified from the RT-cDNA of the developing cell culture and plasmodia, but not amoebae (lanes 6, 7 and 5 respectively: Fig. 3.4). Thus, proP is expressed in the developing cell culture and plasmodia, but not in amoebae. The data agree with the previous northern analysis of proA and proP (Bailey et al., 1999). When proA and proP are amplified from the RT-cDNA samples, they produce intron- bands which shows that the PCR products have been amplified from RT-cDNA molecules and not genomic DNA. Thus, the RT-PCR system seems to be reliable.

The expression of redE was examined by RT-PCR using the same RT-cDNA samples. A 718bp PCR product was amplified from the RT-cDNA of the developing cell culture and plasmodia, but not amoebae (lanes 2, 3 and 1 respectively: Fig. 3.4). Despite the sensitivity of the RT-PCR, redE mRNA could not be detected in amoebae, which suggests redE is not expressed at all in this cell type. The results confirm the northern analysis by showing that redE is expressed in developing cells and plasmodia, but not in amoebae.

3.5 THE GENOMIC MAP OF redE

Southern blotting was used to create a detailed restriction map of redE. The map was needed for the replication studies that are described in Chapter 5 and to identify suitable restriction enzyme(s) for obtaining a genomic clone of redE (section 3.6). The Southern blotting was carried out in two different laboratories; at Leicester University, UK and at the Institut de recherches sur le Cancer, Villejuif, France. The Southern blots made in Leicester used DNA from the haploid strain CL, and were prepared and probed as described in sections 2.4.1 and 2.4.4. While, the Southern blots made in Villejuif used
DNA from the diploid strain M3CIV, and were prepared and probed as described in sections 2.7.1 and 2.7.4. All of the blots were probed with the redE cDNA insert. Representative Southern blots can be seen in Fig. 3.5 and the restriction map of the redE locus can be seen in Fig. 3.6.

The Southern blotting confirmed that redE is a single-copy gene. In the haploid strain, CL, single bands were seen in each lane (e.g. StuI and KpnI: Fig. 3.5) except when the restriction enzyme cut within the cDNA sequence (e.g. Scal: Fig. 3.5; Fig. 3.6). The diploid strain M3CIV, is a derivative of the natural isolate, Wisconsin 1 (Wis 1). It has been shown that Wis 1 and its derivatives frequently have restriction fragment length polymorphisms (RFLPs) between alleles (Schedl and Dove, 1982). In the diploid strain, M3CIV, single bands were seen in each lane (e.g. XbaI BamHI: Fig. 3.5) except when the restriction enzyme cut within the cDNA sequence or there was an allelic polymorphism (e.g. EcoRV: Fig. 3.5).

The genomic map of redE was built in an organised way using several restriction enzyme sites located within the cDNA. There are two EcoRI sites located very close together, approximately 300bp from the 3' end of the cDNA (Fig. 3.6). When a Southern blot containing M3CIV EcoRI digested DNA was probed with the redE cDNA, three bands could be visualised on the blot; there was a faint 2.3kb band and two strong bands of 8kb and 8.8kb (data not shown). It was thought that the 2.3kb band was faint because it stretched downstream of redE and had only 200bp of homology to the cDNA probe, while the 8kb and 8.8kb restriction bands were strong because they stretched upstream of redE and had 750bp of homology to the probe (Fig. 3.6). The two differently sized upstream EcoRI restriction fragments presumably result from an allelic polymorphism.

The restriction enzyme sites located upstream of redE were mapped from the 5' EcoRI site within the cDNA, as described below. When a Southern blot containing EcoRI and XbaI digested DNA was probed with the redE cDNA, two bands were produced; there was a faint 2.3kb band and a strong 1.4kb band (data not shown). As the 8kb and the 8.8kb EcoRI restriction fragments had been cut by XbaI, it was concluded that an XbaI restriction site lay 1.4kb upstream of the 5' EcoRI site within the cDNA (Fig. 3.6). The restriction enzyme sites located downstream of redE were mapped from the XbaI site located just upstream of redE, as described below. When a blot containing XbaI digested DNA was probed for redE a 16kb band was seen (data not shown), which placed an XbaI site 16kb downstream of the site immediately 5' of redE (Fig. 3.6). But when a blot containing XbaI and BamHI digested DNA was probed with the redE cDNA, a band of 4.3kb was visualised on the blot (Fig. 3.5). As the 16kb XbaI fragment had been cut by BamHI, it was concluded that a BamHI site was located 4.3kb downstream of the 5' XbaI site (Fig. 3.6).
3.6 FINDING AND SEQUENCING THE GENOMIC CLONE OF redE

In order to clone the missing 5’ end of redE and some of the upstream promoter sequence a genomic restriction fragment that stretched upstream of the gene was cloned. ML8S was not rescreened for a longer clone of redE because none of the isolated cDNAs have been full length, e.g. frgP, redA and redB (T’Jampens et al., 1997; Bailey et al., 1999). As the redE cDNA is missing a relatively small amount of coding sequence it seemed unlikely that rescreening would yield a longer clone. Restriction mapping showed that an Xbal restriction site was located approximately 1.5kb upstream of the HindIII site that is contained within the 3’ end of the cDNA (Fig. 3.6; position 899-904: Fig. 3.2). This 1.5kb Xbal HindIII restriction fragment seemed an ideal cloning target because it would contain approximately 600bp of novel sequence, including the missing coding sequence. In order to find this restriction fragment a size fractionated genomic DNA library was made using CL genomic DNA.

3.6.1 Constructing a size fractionated genomic library

Genomic DNA was isolated from CL microplasmodia (section 2.2.4). Approximately 60μg of the DNA was digested with Xbal and HindIII and run on a 0.8% agarose gel (sections 2.3.1.3 and 2.3.2). Restriction fragments ranging in size from 1-2kb were isolated from the gel using a Qiagen gel extraction kit (section 2.3.3). The samples were then pooled and phenol extracted before resuspension in a 20μl volume. Meanwhile, 30μg of pBluescript II SK (Stratagene) was digested with Xbal and HindIII, treated with alkaline phosphatase and phenol extracted before resuspension in a 30μl volume (sections 2.3.1.1 and 2.3.4). Three ligations were then set up: (i) 0.2μl of vector was self-ligated (ii) 0.2μl of vector was ligated to 1μl of the size fractionated DNA and (iii) 0.2μl of vector was ligated to 2.5μl of the size fractionated DNA (section 2.3.5). The ligations had a final volume of 10μl. Approximately 3μl of each ligation was transformed into E. coli XL1-Blue and a proportion of each ligation was plated onto LB-amp plates (section 2.3.5). After overnight incubation only two colonies could be seen on the self-ligation plate. In contrast, hundreds could be seen on the other two plates. It was assumed that the ligations had been successful because vector self-ligation occurred at very low frequency. Calculations showed that the ligation which contained 0.2μl of vector and 1μl of the size fractionated DNA had produced the most colonies per μl of insert. Thus, to optimise the size of the genomic library, the remainder of the size fractionated DNA was ligated to the pBluescript in this ratio. After incubation, the ligations were pooled and resuspended in a 21μl volume. The ligation mixture was then transformed into E. coli XL1-Blue in 3μl volumes (section 2.3.5); seven transformations were needed in total. Each of the transformations produced 1ml of transformation mixture. Glycerol was added to each of the mixtures, which were
quick frozen and stored at -80°C (section 2.3.6); the library was called Leicester Library 10 (LL 10). Varying amounts of LL10 were plated onto LB-amp plates and incubated overnight. By counting the number of colonies per plate, it was estimated that the library contained 5.6x10^5 colonies. Twenty colonies were picked at random and the plasmid DNA was extracted from each (section 2.2.6). The DNAs were digested with *SstI* and *XhoI*, and the digests were electrophoresed on a 1% agarose gel (data not shown; sections 2.3.1.2 and 2.3.2). Nineteen of the plasmids contained inserts that ranged in size from 1-2kb. Thus, an estimated 95% of the colonies within LL10 contained inserts within the correct size range.

### 3.6.2 Screening the library for the genomic clone of *redE*

The library was then screened for the 1.5kb *XbaI HindIII* genomic clone of *redE*. Approximately 10μl of the library was plated onto five LB-amp plates, and the plates were incubated overnight. Each plate contained 700-1000 colonies. Colony lifts were carried out (section 2.4.3) and the blots were probed with the *redE* cDNA (section 2.4.4). Several positives came up on the x-ray film, but the plates contained so many colonies that it was impossible to trace each positive back to a single colony. Instead the region of each positive was marked on the plate and then picked and inoculated into 1ml of LB. The number of cells per ml was calculated by counting the bacteria with a haemocytometer. The bacteria were then diluted and an estimated 40 colonies were plated from each 1ml of LB onto a fresh LB-amp plate. The plates were incubated overnight and colony lifts were carried out (section 2.4.3). Several positives came up on x-ray film and this time all of them could be traced back to a single colony. Altogether 11 colonies were picked for analysis and plasmid DNA was extracted from each with a Qiagen miniprep kit (section 2.2.5). Each plasmid was digested with a variety of restriction enzymes, and the digests were electrophoresed on a 1% agarose gel (data not shown; sections 2.3.1.2 and 2.3.2). All of the plasmids had the same digestion pattern indicating they carried identical inserts. The digestion patterns that the 1.5kb *XbaI HindIII* clone would give had been predicted from the cDNA sequence, and the predicted digest patterns were identical to those observed; e.g a *HindIII XbaI* digest released a 1.5kb insert, and a *HindIII NcoI* digest generated two restriction fragments of 3.7kb and 0.8kb. Thus, it seemed likely that all of the plasmids contained the *redE* genomic clone.

### 3.6.3 Sequencing *redE* genomic clone

Of the 11 inserts identified, two were sequenced using the universal reverse and -40 primers (Table 2.3). Automated sequencing reactions were performed as described in section 2.4.5. The sequence from both of the plasmids was identical, confirming they carried identical inserts. When the -40 sequence of one of these plasmids was compared to the *redE* cDNA sequence, there was a 100% match to the 3’ end of the cDNA, showing
that the plasmids contained the redE genomic clone. The reverse primer gave data for approximately 500bp of sequence and the -40 primer gave data for approximately 300bp of sequence; a total of 800bp of sequence had been determined. As the XbaI HindIII clone is approximately 1.5kb in length, around 700bp of sequence was still missing.

To obtain the full sequence of the 1.5kb XbaI HindIII clone, two new sequencing primers were designed that had binding sites in the already sequenced DNA, redE1 and redE2 (positions 151-134 and 794-812 respectively: Fig. 3.7; Table 2.3). The redE genomic clone was sequenced with these primers (section 2.4.5); the redE1 primer gave approximately 500bp of sequence, and the redE2 primer gave for approximately 450bp of sequence. When the redE1 and the redE2 sequences were compared, there was a 63bp region of overlap showing that the complete sequence of the genomic clone had been determined. The fragment was 1533bp in length and overlapped with the cDNA for 900bp (EMBL accession number AJ297388). Thus, redE does not contain introns within this region. The genomic clone and the cDNA sequences were combined to give Fig. 3.7.

### 3.7 ANALYSIS OF THE redE GENE

The northern analysis showed that the redE cDNA clone is missing either 50bp or 200bp (section 3.4.1). The XbaI HindIII genomic clone should contain the start of the redE coding sequence as it contains 600bp of new sequence stretching upstream of the start of the cDNA sequence. In order to try and construct an ORF that started in the genomic clone and read through to the cDNA clone, putative start codons (ATG) were identified in the genomic sequence. But an ORF that read from the start codons to the beginning of the cDNA sequence could not be identified because every start codon was followed by a stop codon in the same reading frame. It was hypothesised that an intron, located between the start codon and the start of the cDNA sequence, was responsible for disrupting the ORF.

The intron boundaries of nuclear mRNAs have conserved sequences; the dinucleotide, GT is always at the 5' intron boundary and the dinucleotide, AG is always at the 3' intron boundary (Breathnach and Chambon, 1981). All of the P. polycephalum genes studied so far have these sequences at their splice junctions, e.g. arda, proA, proP, redA and redB (Nader et al., 1986; Binette et al., 1990; Bailey et al., 1999). This GT-AG rule was used to identify the position of the putative intron in the redE genomic sequence. As the cDNA clone is missing a relatively small amount of sequence, the start codons that were closest to the start of the cDNA sequence were analysed first. The first start codon was 85bp upstream of the start of the cDNA sequence (atg position 58: Fig. 3.7), the second start codon was 99bp upstream of the start of the cDNA sequence (atg position 44: Fig. 3.7) and the third start codon was 142bp upstream of the start of the cDNA sequence (ATG position 1: Fig. 3.7). Each of these start codons was followed by a stop codon in the same
reading frame, (tag and tag: positions 64 and 107: Fig. 3.7). The first step was to identify every GT and AG dinucleotide located between the outer start codon and the beginning of the cDNA sequence. Then the region between every combination of GT and AG dinucleotides was deleted. This mimicked removing an intron from the genomic sequence. Each sequence was then checked to see whether an uninterrupted ORF, that read from an ATG and through the cDNA sequence had been created.

Only one deletion created an uninterrupted ORF which stretched from the outermost ATG to the stop codon in the cDNA sequence (position 1-1072: Fig. 3.7). The deleted region, and thus the putative intron was 114bp in length (position 22-135: Fig. 3.7). This gave an extra 28bp of protein coding sequence; 21bp between the putative start codon and the 5' end of the putative intron (position 1-21: Fig. 3.7), and 7bp between the 3' end of the intron and the start of the cDNA sequence (position 136-142: Fig. 3.7). Thus, the putative coding region is 957bp in length (Fig. 3.7). Further experiments are needed to determine whether the putative start codon and intron have been positioned correctly. This is discussed further in section 3.9. The amino acid sequence of the RedE protein was deduced from the putative coding region (Fig. 3.7). The putative RedE protein contains 319 aa's. The analysis of this sequence is discussed in section 3.8.

The redE cDNA clone contains 152bp of 3’ untranslated region (Fig. 3.7). The consensus polyadenylation signal AATAAA (Nader et al., 1986; Morita, 1998; Bailey et al., 1999) or close variants of this, such as TATAAA, ATTTAA, AATATA or AAATAA (Hamelin et al., 1988; Binette et al., 1990; T’Jampens et al., 1997 and 1999) are generally found in the 3’ untranslated regions of P. Polycephalum genes. When the 3’ untranslated sequence of redE was analysed, it became clear that it did not contain any of these polyadenylation signals. A study by van Helden et al. (2000), involving statistical analysis of sequences located downstream of ORFs in yeast, led to the identification of many putative polyadenylation signals. Three of these putative signals, ATGTAC, TGTACA and TAAGTA were found in the 3’ untranslated region of redE and could possibly represent polyadenylation signals. The sequence (atgtaca) contains the first two signals; ATGTAC is underlined and TGTACA is shown in bold (position 1166: Fig. 3.7). The third signal, (taagta) is repeated twice (at positions 1198 and 1208: Fig. 3.7).

The redE cDNA hybridised to two mRNAs on a northern blot (section 3.4.1). It is not known whether the redE cDNA was reverse transcribed from the larger or the smaller mRNA. As they are approximately 1300nt and 1150nt in length, the 1102bp cDNA is missing either 200bp or 50bp respectively. As redE is a single-copy gene (sections 3.2 and 3.5), both of the redE mRNAs must arise from this gene. It is unlikely that the two mRNAs are generated by differential splicing because redE has only one intron. Instead,
the two mRNAs probably have identical coding regions, but differ in the amount of untranslated sequence they carry.

There are several ways by which this can occur:

(i) redE could have two transcription start sites. This would produce two transcripts that carried different lengths of 5' untranslated sequence.

(ii) redE could have multiple polyadenylation signals. This would produce transcripts with different lengths of 3' untranslated region.

(iii) The two redE mRNAs could have differently sized poly(A)+ tails.

3.8 ANALYSIS OF THE RedE PROTEIN

The RedE protein sequence was analysed using the Wisconsin Package, Genetics Computer Group (Madison, USA) on the Leicester University mainframe computer. The RedE protein contains 319 aa's (section 3.7), of which 8% are acidic and 16% are basic. The protein is rich in Proline (15%) and Serine (13%) residues (Fig. 3.7) and has a predicted molecular weight of 36kDa. A hydrophilicity plot showed that RedE is a hydrophilic protein (data not shown). Secondary structure calculations (data not shown) showed that most regions of RedE will not adopt an α-helical or a β-pleated sheet structure. Several sites for potential protein modification were found within the deduced RedE protein sequence; there were three potential N-linked glycosylation sites, five potential casein kinase II phosphorylation sites and five potential protein kinase C phosphorylation sites (Fig. 3.8). These three types of protein modification are now discussed in turn.

N-linked glycosylation occurs when a chain of sugar molecules is covalently linked to the NH₂ group of an asparagine residue within the aa consensus sequence Asn-x-Ser/Thr (N-x-S/T), where x can be any aa except proline (Bause, 1983). The potential N-glycosylation sites (NMS), (NDS) and (NGS) were found within RedE protein sequence (positions 53, 158 and 278 respectively: Fig. 3.8). The occurrence of potential glycosylation sites within the RedE protein sequence does not give clues to the function of RedE because a range of proteins that perform a variety of functions are glycosylated proteins or 'glycoproteins' e.g. human plasma protein C (a serine protease: Miletich and Broze, 1990), ovalbumin (a secretory protein: Pless and Lennarz, 1977) and human IgG (Mathews and van Holde, 1990). Most of the proteins that are sequestered in the endoplasmic reticulum (ER) are glycoproteins that are later transported to the golgi apparatus, the plasma membrane, the extracellular space, or to lysosomes. Cytosolic proteins are never glycosylated in this way because the enzyme that catalyses the addition of the complex sugar chain is only found in the ER (Alberts et al., 1989). The occurrence of potential N-linked glycosylation sites within the aa sequence of RedE does not necessarily mean that the protein is glycosylated.
because it may be a cytosolic protein and two of the three potential glycosylation sites have a proline residue C-terminal to the serine/threonine residue which significantly decreases the occurrence of glycosylation (Bause, 1983).

The casein kinase II (CK II) family of enzymes catalyse the addition of a phosphate group to a Ser/Thr residue within the aa consensus sequence Ser/Thr-x-x-Asp/Glu (S/T-x-x-D/E), where x can be any aa (reviewed in Pinna, 1990). The potential CK II phosphorylation sites (TPLD), (TRED), (TVHD), (TIRE) and (SGGD) were found within the RedE protein sequence (positions 93, 108, 249, 256 and 282 respectively: Fig. 3.8).

Protein kinase C (PK C) enzymes catalyse the addition of a phosphate group to a Ser/Thr residue within the aa consensus sequence Ser/Thr-x-Arg/Lys (S/T-x-R/K), where x can be any aa (Kishimoto et al., 1985). The potential PK C phosphorylation sites (SPK), (TAR), (SWR), (SSR) and (TIR) were found within the RedE protein sequence (positions 8, 19, 84, 150 and 256: Fig. 3.8). The occurrence of potential phosphorylation sites within RedE sequence does not give clues to the function of the protein because a range of proteins that perform a variety of functions are phosphorylated by casein kinase II-type enzymes and protein kinase C, e.g. the yeast ribosomal stalk proteins (Ballesta et al., 1999), human NAP-1 and NAP-2 (histone chaperones: Rodriguez et al., 2000), P. polycephalum FrgP (an actin binding protein: de Corte et al., 1996) and human moesin (an actin cytoskeletal-membrane linkage protein: Pietromonaco et al., 1998).

BLAST, FASTA, and TFASTA database searches showed that RedE was significantly homologous to a variety of proteins that are also rich in proline and serine residues. A closer examination of the protein alignments showed that RedE did not share any real homology with these proteins because the compared proteins did not have common motifs, the alignments often contained gaps and the identical amino acids were almost always proline or serine residues. For example, RedE is homologous to the hydroxyproline-rich glycoproteins (HRGP’s) of Sorghum vulgare (Raz et al., 1991) and Zea mays (Stiefel et al., 1988). The protein is 23.5% identical and a further 16.3% similar to the S. vulgare HRGP over a 251aa region, and 24% identical and a further 19% similar to the Z. mays HRGP over a 238aa region (data not shown). HRGP’s constitute a family of glycoproteins that are structural components of plant primary cell walls. These proteins are rich in hydroxyproline/proline, basic aa’s, serine and tyrosine (Varner and Lin, 1989). The HRGP protein sequences contain repeated peptides of five and six aa’s (Raz et al., 1991; Stiefel et al., 1988). The maize polypeptide chain contains hydrophobic segments interspersed with hydrophilic segments (Stiefel et al., 1988). Although RedE and the HRGP’s are similar in amino acid content, it is unlikely that RedE is related to the HRGP’s because it does not have any of the other features that are associated with these proteins, such as repeated
polypeptides and hydrophobic segments. In addition, *P. polycephalum* cells do not have a plant primary cell wall.

3.9 DISCUSSION

I screened some of the 150 frozen stocks that were identified in the first round screen of the subtracted cDNA library (ML8S) and identified six colonies that contained plasmids with inserts. The expression of the six cDNAs was then analysed by northern blotting. Two of the cDNAs, D11/1100 and A18/1020 were identified as potential red genes. The analysis of D11/1100 was the focus of this chapter. The D11/1100 cDNA clone was fully sequenced. The 1102bp partial cDNA clone contained 929bp of coding sequence, 152bp of 3' untranslated sequence and 18bp of poly(A)+ tail. This is consistent with studies of other *P. polycephalum* genes which have identified 3' untranslated regions ranging from 12-195bp (Hamelin *et al*., 1988; Binette *et al*., 1990; Bailey *et al*., 1999). A clear polyadenylation signal could not be identified in the 3' untranslated region of redE.

Southern blotting analysis showed that D11/1100 represents a single-copy gene. The expression of the D11/1100 gene was analysed by northern blotting analysis and RT-PCR. The northern blotting showed that the D11/1100 represents a red gene; the gene is not expressed in amoebae, is expressed at high levels in developing cells and is expressed at low levels in plasmodia. Thus, D11/1100 was renamed redE. The RT-PCR confirmed the results of the northern blotting analysis by showing that the redE mRNA is only present in the RNA of developing cells and plasmodia.

The northern blotting showed that redE has two mRNAs of 1300nt and 1150nt indicating the 1102bp cDNA clone is missing a relatively small amount of coding sequence. It was suggested that the two redE mRNAs arise because the gene has multiple transcription start sites and/or polyadenylation signals, or the two mRNAs could simply have different lengths of poly(A)+ tail. A similar situation was observed for the profilin gene, proA. The proA cDNA hybridises to two mRNAs of 500nt and 600nt on a northern blot, despite being a single-copy gene (Binette *et al*., 1990). Primer extension experiments showed that proA has two transcription start sites, and sequencing 12 different proA cDNAs showed that proA has two polyadenylation signals; proA may have as many as four different types of mRNA. It appears that proA hybridises to only two mRNAs because some of the transcripts are so close in size (Binette *et al*., 1990). Similar experiments would need to be carried out to see if redE has two transcription start or polyadenylation sites. Several possible approaches are outlined below.

S₁ nuclease mapping or primer extension could be used to map the transcription start site(s) of redE (Binette *et al*., 1990; Mathews and van Holde, 1990). In S₁ nuclease
mapping, the redE genomic clone would be 5' end-labelled and cleaved asymmetrically so that only the antisense strand was labelled. The DNA would then be denatured and hybridised to an RNA sample containing the redE transcript. After hybridisation, the single stranded DNA and RNA is removed from the sample by S1 nuclease digestion. The digestion leaves the DNA-RNA hybrid of interest and the size of the hybrid is determined by running the sample on a sequencing gel. The size of the hybrid reveals the position of the transcription start site(s), because the hybrid stretches from the 5' end of the antisense strand of the genomic clone to the 5' end of the transcript of interest (Mathews and van Holde, 1990). In primer extension, the redE cDNA clone would be digested to give a smaller restriction fragment that is located close to the start of the cDNA sequence. The fragment would then be end-labelled, denatured and hybridised to an RNA sample containing the redE transcript. This fragment then primes a reverse transcriptase reaction which extends the 3' end of the DNA primer until it reaches the 5' end of the redE transcript, and the size of the extended product is determined by running the sample on a sequencing gel. The length of the extended product reveals the position of the transcription start site(s) because the product size is the distance between the 5' end of the primer and the 5' end of the redE transcript (Binette et al., 1990; Mathews and van Holde, 1990).

The subtracted cDNA library, ML8S, could be rescreened for more redE cDNA clones to determine whether redE uses different polyadenylation signals. A proportion of the subtracted cDNA library would be plated onto several LB-amp plates and the plates would be incubated. Colony blots would be carried out and the blots would be probed with the redE cDNA. Any positive colonies could be isolated and the cells would be grown for plasmid isolation. Sequencing the plasmids would reveal whether any of the clones have different lengths of 3' untranslated region. RNaseH digestion could be used to see if the two redE mRNAs have different lengths of poly(A)+ tail (Binette et al., 1990). In this procedure, total RNA from the developing culture in which 56% of the cells were committed to development would be hybridised to oligo(dT) and the poly(A)+ tails would be removed by RNase H digestion. The protected fragments would then be subjected to the northern blotting procedure and the resulting blot would be probed with the redE cDNA. If the two mRNAs have differently sized poly(A)+ tails, a single mRNA would be visualised on the blot. In contrast, if the two mRNAs have the same lengths of poly(A)+ tail, two mRNAs would be visualised on the blot. These transcripts would be smaller than the previously observed 1300nt and 1150nt transcripts, because they would be missing their poly(A)+ tails.

In order to obtain the missing 5' end of redE coding region and some of the upstream promoter sequence, a genomic restriction fragment that stretched upstream of redE was cloned. The genomic clone was fully sequenced. The 1533bp genomic clone contained
approximately 600bp of new upstream sequence. Analysis of the genomic clone led to the identification of a putative start codon and a putative intron of 114bp. This is within the intron size range observed for other *P. polycephalum* genes; introns ranging from 58-278bp have been identified (Nader *et al.*, 1986; Binette *et al.*, 1990; Bailey *et al.*, 1999). S; nuclease mapping or primer extension experiments would determine whether the putative start codon and intron have been positioned correctly, in addition to revealing whether redE has two transcription start sites (see above). The putative redE coding region is 957bp in length. The aa sequence of the RedE protein sequence was deduced from the putative coding region. The protein contains 319aa’s and is rich in proline and serine residues.

The RedE protein sequence was analysed using the Wisconsin Package, Genetics Computer Group (Madison, USA). These analyses showed that RedE contains three potential N-linked glycosylation sites, five potential casein kinase II phosphorylation sites and five potential protein kinase C phosphorylation sites. The following experiments could be carried out as a first step towards determining whether RedE is phosphorylated. The first step would be to purify the RedE protein. A RedE recombinant protein could be made by cloning the putative redE coding region into a vector containing, for example, the glutathione S-transferase gene (gst). The putative redE coding region would be cloned immediately downstream of gst so that gst-redE forms a continuous reading frame that is under the control of a single promoter. The vector would then be transformed into *E. coli* and the expression of the GST/RedE fusion protein would be induced. The protein would be isolated from cell extracts using a glutathione-sepharose column. The recombinant protein could be phosphorylated *in vitro* by casein kinase II or protein kinase C using 32P labelled dATP, and the sample would be subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) to see if the protein is labelled, and was therefore phosphorylated. Alternatively, mass spectrometry could be used to determine whether the fusion protein had a higher molecular weight after the *in vitro* phosphorylation, and was therefore phosphorylated.

Database searches showed that although RedE was significantly homologous to a variety of proteins that are also rich in proline and serine residues, it did not share any real homology with these proteins. Thus, the function of RedE is unknown. Some possible approaches to investigating the function of RedE are outlined below. Immunofluorescence microscopy could be used to localise the RedE protein. The immunolocalisation could be done by cloning redE into a recombinant protein vector that contained a tag such as gst (see above) or gfp (green fluorescent protein), an intrinsically fluorescent protein. The redE coding region would be cloned so that gst-redE or gfp-redE formed a continuous
reading frame that was under the control of a constitutive promoter. The constructs would then be transiently transformed into cells and the redE-gst fusion protein would be localised using Anti-GST antibodies and the redE-gfp fusion protein would be localised by visualising the cells under a certain wavelength of light. The gfp system has been successfully used in a range of organisms, e.g. Bacillus substilis (Lewis and Errington, 1996) and the slime mould Polyphondylium (Fey and Cox, 1997). Alternatively, the immunolocalisation could be carried out using anti-RedE antibodies that were raised against purified recombinant protein (see above). Such experiments would localise RedE in addition to pinpointing the expression of the gene to specific cells. The redE-gfp system could be adapted to make it particularly suitable for use in P. polycephalum. The first step would be to make a construct containing a selectable marker, such as hph, and redE-gfp under the control of the redE promoter. The construct would then be transformed into amoebae and stable transformants would be selected on the basis of Hyg resistance. In stable transformants, the expression of redE and the localisation of the protein could be examined in individual cells throughout development. This type of system has been successfully used in Drosophila melanogaster (Yeh et al., 1995).

The red genes are ideal candidates for gene knockout because they are not expressed in amoebae and amoebal growth in knockout strains would be unaffected. Gene knockout by homologous integration has previously been achieved for the actinD gene, ardD (Burland and Pallotta, 1995). Several constructs containing hph and a mutated copy of the redE coding region could be transformed into amoebae. Stable transformants would be selected on the basis of hygromycin resistance and the phenotype of any mutant strains with homologous integration would be analysed by time-lapse filming and immunofluorescence microscopy. Alternatively, gene knockout could be achieved using antisense experiments. A construct which contained hph and a reversed copy of the redE coding region would be transformed into amoebae and stable transformants would be selected on the basis of Hyg resistance. The phenotype of any stable transformants would then be analysed as above. The integrated construct would knockout the function of wild-type redE because the reversed copy of redE would create an antisense transcript that forms a duplex with the wild-type redE mRNA, thus preventing its translation (reviewed in Vanhée-Brossollet and Vanquero, 1998). The function of redE could also be investigated by mis-expressing the gene. Several constructs containing hph and the redE coding region under the control of various promoters (e.g. the promoter of a constitutively expressed gene such as PardC) could be transformed into amoebae and stable transformants would be selected on the basis of Hyg resistance. The phenotype of any transformants would then be analysed as above. The work in chapter 6 describes the development of such a system.
(a) purified D11/1100 plasmid was sequenced with the universal reverse and -40 sequencing primers (Table 2.3) as described in section 2.4.5. The sequence that was read by the reverse and -40 primers is shaded.

(b) subclone 1 was generated from the D11/1100 cDNA as described in section 3.3.1. The subclone sequence that was read by the reverse and -40 primers is shaded. The hatched region indicates sequence overlap between the reverse and the -40 sequences. The D11/1100 cDNA and the subclone 1 sequences were aligned using the Wisconsin Package, Genetics Computer Group (Madison, USA). The comparisons showed that subclone 1 is the 636bp Sau3AI restriction fragment of D11/1100, as indicated by the dashed lines.
Figure 3.2: The nucleotide and deduced amino acid sequence of the D11/1100 cDNA clone

The nucleotide sequence of the 1102bp cDNA clone (EMBL accession number AJ297387) was determined by sequencing the D11/1100 cDNA and subclone 1 (section 3.3; Figure 3.1). The coding region is shown in uppercase letters and the non-coding region is shown in lowercase letters. The open reading frame starts with the third base of the cDNA, and ends with a stop codon at base 930 (TAG), indicated by * in the amino acid sequence. The coding region is followed by 152bp of 3' untranslated sequence and 18bp of poly(A)+ tail. The nucleotide and protein sequences are partial, as cDNA does not begin with an ATG start codon. The deduced amino acid sequence is shown in standard single-letter code.

The positions of two PCR primers, redEd2 and redEm2 are marked on the sequence at bases 215-234 and 913-932 respectively (see section 3.4.2). The position of the HindIII restriction site is marked on the sequence at bases 899-904 (see section 3.6).
1 AA AAA GAA TTA AGT AAG GAA ATC TAC ACA GCT CGG AGA TCA CAA CAG TTC
   R E L S K E I Y T A R R S Q Q F
51 CAG CCA CAT CAA ACA GAA CCA AGC ATT GGA CAA AGA CCT TCT CTT TAC AAC
   Q P H Q T E P S I G Q R P S Y N
102 CTC CCC CTT TCC TCT CAC AAT TCT CAC AAT ATG TCG CCA TGG AGA CCA CCC
   L P S S H N S H N M S P W R P P
153 AAG GAC CAC GAA GAT CCA CCA TCA CAA GTA CTA GGG TGT AGT ACT TCC TAT CTC
   K D H E D P S P V L G C S T S Y P
204 GCC CCA CTG AGA ATG ACA TCC TGG AGG CCA GAT AGA GAA GAG GAA ACG CCA
   redEd2
   A P L R M T S W R P V R E R E T P
255 CTA GAT TTT CCC CCG CAC CCA CTT CCA TGC AAA CCC TAC AGC AGA GAA GAT
   L D F P P P C K P Y T R E D
306 AGC CAC CAA ATA TCT GGT AAT CCT CAT ACC AAA ACA ACA CTA CCT CCT CTT
   S H Q I S G N P H T K T T L P L
357 CTC ATT GAC GCC CAG CAT TAC TCA CAA CAA AAC AAC AAC CAA TAC GCA CAA
   L I D A Q H Y S Q Q N N Q Y A Q
408 AAA CCA TTC CTA TCT TCC CAC AAT TCT CAC AAT ATG TCG CCA TGG AGA CCA
   redEm2
   K P F L S S S R S L Y P R N D S P S
459 AGA GAA CTT CCA CAC ATC CGC GAG TCA CTT CCA AAT AAT ATC TGG TCG AAG
   R E L P H I R E L P N N I W S K
500 ATT TCC CCA CCT AAA CAA AAC AAG TGC CCT TCT CCA GAT AAC TTC GCG CCA
   I S P P K Q N K P C S P D N F R P
561 GAC GCT AAG TCC TGC TAC ATT TGC GAA GGT CTG CAT CAC TCC CTA CCA TTT CAA
   D A N S C Y I C E G L H H F P F Q
612 TTT CCC TTC CTC CAT TTT CCC CTC CAG AAC CCA CCA CCT GCC ACT TCC TTC
   F P F L H L P L Q N P P A T L F
663 CGC ACA TAC GCA CAC CTC CAC CGT ATT CAC CAC CGC CAC CCC TCG CTT CTA CTT
   C R T Y A H L H V H L R H P S L L P
714 CCC ACC GTC CAC GAC GAG CAA ACG ACT ATA CGG GAG AAT TCA AGA GAG AAT
   P T V H D E Q T T I R E N S R E N
765 TTT GTG CCC TTT TGG CCA AGC ATG ACG AAG CGC GAT CAC AAT GGG AGC ACT
   F V P F W P S M Q K R D H N G S T
816 TCT GGA GGA GAC GCG AGT GCA CAG GGA GAC GCC CTG GCA TCA AGT CAT ACC
   S G G D A S A Q G D G R A S S H T
867 TTG CAT TTT GTG ACC TGT ACC CTG AAT TCA CCA AGC TTA GCC GTC TCC ACA
   HindIII
   L H F V T C T L N S P S L G V P T
918 AGT ACC GTG TCC TAG gcactcgccaaaatggtagtgcatctcatcaagcctccaacagc
   redEm2
   S T V S *
981 ttaccttgtttttgcctgtttatgtgtatagtattcattacggtctcatgtcgtaactaatgagccaaacc
1048 actgcctgtaatagtaagtagataagagttttagtaatatttaaaaaaaaaaaaaaaaaaaa 1102
   poly (A)+ tail
Total RNA was isolated from CL amoebae (A), microplasmodia (Mi), macroplasmodia (Ma) and from a variety of developing cultures (1-56%; Table 2.1) as described in section 2.2.2. 20μg of each total RNA sample, and 3μg of RNA size markers (Promega) were denatured and run on a 1.1% agarose gel as described in section 2.4.2. The RNA was transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.2).

The cDNA clones for (a) D11/1100, (b) proP (Binette et al., 1990) and (c) actin (Hamelin et al., 1988) were radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised sequentially at high stringency (65°C) to the same northern blot. The blot was exposed to x-ray film at -80°C with an intensifying screen. D11/1100 hybridised to 1300nt and 1150nt mRNAs (12 day exposure), proP hybridised to a 520nt mRNA (1 day exposure) and actin hybridised to 1400nt mRNA (2 hour exposure).
Figure 3.4: RT-PCR analysis of the expression of D11/1100

Total RNA was isolated from CL amoebae, a developing CL culture in which 56% of the cells were committed to development (Table 2.1) and from LU352 microplasmodia as described in sections 2.2.2 and 2.5.4. 10μg of each total RNA sample was reverse transcribed to make an RT-cDNA sample (section 2.5.4). A 1μl volume of each RT-cDNA sample was amplified by PCR (section 2.5.4) with primers for either D11/1100, or proA and proP (Table 2.3). 100ng of M3CIV genomic DNA was amplified by PCR with primers for either proA or proP. 10μl of each PCR reaction was run on a 2% agarose gel (section 2.6.2) with φX174 HaeIII DNA size markers. The gel was stained with ethidium bromide and photographed under UV illumination.

Lane 1: Amoebal RT-cDNA + D11/1100 primers.
Lane 2: Developing culture RT-cDNA + D11/1100 primers.
Lane 3: Plasmodial RT-cDNA + D11/1100 primers.
Lane 4: Genomic DNA + proA primers.
Lane 5: Amoebal RT-cDNA + proA and proP primers.
Lane 6: Developing culture RT-cDNA + proA and proP primers.
Lane 7: Plasmodial RT-cDNA + proA and proP primers.
Lane 8: Genomic DNA + proP primers.
Figure 3.5: Southern blotting analysis of redE

(a) 3μg of genomic DNA from CL microplasmodia was digested with the restriction enzymes shown (section 2.3.1.3) and run on a 0.8% agarose gel with λ HindIII DNA size markers (section 2.3.2). The DNA in the gel was denatured, transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.1). The redE cDNA clone was radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised at high stringency (65°C) to the southern blot. The blot was exposed to x-ray film at -80°C with an intensifying screen for approximately 1 week.

(b) 3μg of genomic DNA from M3CIV macroplasmodia was digested with the restriction enzymes shown (section 2.6.1.2) and run on a 0.8% agarose gel with λ HindIII DNA size markers (section 2.6.2). The DNA was denatured, transferred to membrane using appligene vacuum blotting apparatus and fixed to the membrane by baking at 80°C for 1 hour (section 2.7.1). The redE cDNA clone was radio-labelled using a NEN kit (section 2.7.4) and hybridised at high stringency (65°C) to the southern blot. The blot was exposed to x-ray film at -80°C with two intensifying screens for approximately 1 week. The relative migration of the λ HindIII DNA size markers is shown in bp.
Figure 3.6: Restriction map of the \textit{redE} locus

Most of the restriction sites were mapped in the diploid strain, M3CIV. Restriction sites that were mapped in the haploid strain, CL are underlined. Allelic polymorphisms are marked *. The restriction sites located 5' of \textit{redE} were mapped from the most 5' internal EcoRI restriction site, which is indicated in bold. The restriction sites located 3' of \textit{redE} were mapped from the 5' XbaI restriction site, which is also indicated in bold. The sequential order of restriction sites which are close together may not be correct because the map was constructed with the data obtained from several southern blots.
The nucleotide sequence of the 1533bp redE genomic clone (EMBL accession number AJ297388) was determined by sequencing with the reverse, -40, redE1 and redE2 sequencing primers (Table 2.3) as described in section 2.4.5. The sequences were aligned using the Wisconsin Package, Genetics Computer Group (Madison, USA). The positions of the sequencing primers, redE1 and redE2 are marked on the sequence at bases 151-134 and 794-812 respectively. The nucleotide sequences of the genomic and the cDNA (Figure 3.2; EMBL accession number AJ297387) clones were combined to give this figure. The genomic clone starts at the XbaI site at base -487 and ends at the HindIII site at base 1046. The cDNA clone starts at base 143, marked *, and ends at base 1244 with the poly(A)+ tail.

The two putative ATG start codons are marked in bold at bases 1 and 44 (ATG). The stop codons that follow these start codons are underlined at bases 64 and 107. The deduced coding region is shown in uppercase letters, and the non-coding region is shown in lowercase letters. The deduced intron starts with the conserved 5' intron splice site, (gt), at base 22 and ends with the conserved 3' intron splice site, (ag), at base 135. The coding region starts with the ATG at base 1 (ATG) and ends with a TAG stop codon at base 1072 (TAG), indicated by * in the amino acid sequence. The deduced amino acid sequence is shown in standard single-letter code. Four putative polyadenylation signals were identified in the 3' untranslated region. The sequence ATGTACA at base 1166 contains two of the signals; ATGTAC is underlined and TGTACA is marked in bold (atgtaca). The other signal, TAAGTA, is repeated twice at bases 1198 and 1208 (taagta).
TCA AGT CAT ACC TTG CAT TTT GTG ACC TGT ACC CTG AAT TCA CCA AGC TTA
SSHTLHVFVTCFTLNSPSL

HindIII

1048 GGC GTA CCC ACA AGT ACC GTG TCC TAGgcactgcacaaattgtgagtgcagttatatcaag
GVPTSSTVS *

1007 gcccactccccacagctttacctctctctgttttcgccgtattgtgatgtgtaatcggttcaatgtacca

1174 ctattgaggccaaaccactgctgttaagttaaggtaaagttaaagtttaatattaaaaaaaaaaaa
poly (A)+ tail

1241 aaaa 1244
Figure 3.8: Potential sites for protein modification within the amino acid sequence of RedE

The amino acid sequence of RedE contains several potential sites for protein modification. Three N-linked glycosylation sites are shown in bold; (NMS), (NDS) and (NGS) at bases 53, 158 and 278 respectively. Five casein kinase II phosphorylation sites are underlined; (TPLD), (TRED), (TVHD), (TIRE) and (SGGD) at bases 93, 108, 249, 256 and 282. Finally, five protein kinase C phosphorylation sites are shown in bold and underlined; (SPK), (TAR), (SWR), (SSR) and (TIR) at bases 8, 19, 84, 150 and 256. These consensus sequences were identified with the Wisconsin Package, Genetics Computer Group (Madison, USA).
CHAPTER 4: THE ANALYSIS OF *mynD*

4.1 INTRODUCTION

In the first round screen of the subtracted cDNA library, ML8S, 150 colonies of interest were identified (Bailey *et al.*, 1999; section 1.3.3.2). I screened some of the 150 frozen stocks and identified six colonies that contained plasmid with a single insert (section 3.2). Northern blotting analysis suggested that two of the six cDNAs represented *red* genes, and these clones were chosen for further analysis (section 3.2). The previous Chapter described the analysis of the first clone, D11/1100 and this Chapter describes the analysis of the second, A18/1020.

The work described in this Chapter is published in Bailey *et al.* (1999). I acknowledge two undergraduate project students, Lorraine King and Aki Kobayashi for their contribution to this work and Laszlo Nyitray (Eotvos University) for sharing unpublished findings.

RESULTS

4.2 THE EXPRESSION OF A18/1020 DURING THE AMOEBA-L-PLASMODIAL TRANSITION

The preliminary northern analysis suggested that A18/1020 represented a *red* gene (section 3.2). In order to verify this, a more detailed analysis of its expression was carried out by northern blotting using more RNA samples. The expression of A18/1020 was analysed during both apogamic and sexual development.

4.2.1 The expression of A18/1020 during apogamic development

All of the RNA samples were from the apogamic strain, CL. A northern blot was made (section 2.4.2) which contained 10μg of total RNA from (i) amoebae (ii) microplasmodia (iii) macroplasmodia and (iv) a variety of CL developing cultures, each of which contained a different proportion of cells that were committed to apogamic development, ranging from 1-56% (1-56%: Table 2.1; Bailey *et al.*, 1999; section 2.2.2). The developing culture RNA samples that were used to characterise the expression of *red A*, *redB* and *redE* (sections 1.3.3.3 and 3.4.1) were used in this study, so that the expression of A18/1020 could be directly compared to that of the other *red* genes. The blot was probed at high stringency (65°C; section 2.4.4) with the cDNA clones for A18/1020, *proP* (LAV1-5: Binette *et al.*, 1990) and actin (pPpA35: Hamelin *et al.*, 1988; Fig. 4.1). The probing was repeated several times using different northern blots, and the results shown in Fig. 4.1 are representative of those obtained.
A18/1020 hybridised to a single mRNA of approximately 6000nt (Fig. 4.1a). As the cDNA clone is 1020bp in length, it is missing approximately 5kb. The A18/1020 mRNA could not be detected in the RNA from amoebae or the developing culture in which 1% of the cells were committed, but was detected in the RNA from all of the other cell cultures (Fig. 4.1a). The transcript levels increased with the percentage of committed cells to reach a maximum in the 56% sample (Fig. 4.1a). The message was detected at slightly lower levels in the RNA from microplasmodia, and barely at all in the RNA from macroplasmodia (Fig. 4.1a). The expression pattern of A18/1020 is very similar to that of redA, redB (section 1.3.3.3) and redE (section 3.4.1). Like these genes, A18/1020 is expressed primarily during development.

Some control experiments were carried out before A18/1020 was finally designated a red gene. The blot was probed for actin and proP to allow the expression pattern of a plasmodium-specific and a constitutively expressed gene to be directly compared to that of A18/1020. As expected, actin and proP hybridised to 1400nt and 520nt mRNAs respectively (Pallotta et al., 1986; Binette et al., 1990). All of the RNA samples contain easily detectable actin transcript which is not significantly degraded (Fig. 4.1c). As expected, the proP message was not detected in the RNA from amoebae or the developing cell culture in which 1% of the cells were committed to development (Fig. 4.1b; Bailey et al., 1999). The proP message was detected in the RNA from all of the other developing cultures, and the amount of proP mRNA increased with the percentage of committed cells (Fig. 4.1b). The RNA from microplasmodia and macroplasmodia contained the most proP mRNA (Fig. 4.1b). The actin control shows that the blot contains intact amoebal RNA. As the expression of A18/1020 cannot be detected in this sample, the cDNA cannot represent a constitutively expressed or amoeba-specific gene. The proP control shows that A18/1020 does not represent a plasmodium-specific gene, as the expression of A18/1020 is clearly different to that of proP. Thus, A18/1020 does not represent a constitutively expressed, amoeba-specific or plasmodium-specific gene. As the expression pattern of A18/1020 is the same as that of redA, redB and redE, A18/1020 was designated a red gene. However, A18/1020 was not renamed redF for reasons that become clear later.

4.2.2 The expression of A18/1020 during sexual development

To see how far the similarity between apogamic and heterothallic development extended, Bailey et al. (1999) examined the expression of actin, proA, proP, redA and redB during heterothallic development by northern blotting, using RNA that was isolated from heterothallic cultures which contained a high proportion of fusion cells and zygotes (Barber, 1998). A brief description of how these RNA samples were obtained is given below. Using 2-strain matings with amoebae that were heteroallelic at the mat loci, Barber (1998) obtained cultures in which 2%, 4%, 17% and 24% of the cells were committed to
development and the large majority of these cells were fusion cells and zygotes (Table 2.2). But Barber (1998) found that in cultures in which more than 25% of the cells were committed to development, an unsatisfactory proportion of these cells were multinucleate cells presumably because mating occurred with poor synchrony. Barber (1998) achieved improved mating synchrony in 10-strain matings because an amoeba in a 10-strain mating is compatible with 90% of the amoebae, whereas an amoeba in a 2-strain mating is compatible with only 50% of the amoebae. Thus, using 10-strain matings with amoebae that were heteroallelic at the mat loci, Barber (1998) obtained cultures in which 38% and 41% of the cells were committed to development and the large majority of these cells were fusion cells and zygotes (Table 2.2). The northern analysis showed that actin, proA, proP, redA and redB had the same basic patterns of expression in both heterothallic and apogamic development (Bailey et al., 1999).

The northern blot analysis of A18/1020 was extended by examining the expression of the gene during heterothallic development. A northern blot was made (section 2.4.2) which contained 10μg of total RNA from (i) CH508 amoebae (ii) LU648 amoebae (iii) microplasmodia formed by crossing CH508 and LU648 amoebae (iv) macroplasmodia formed by crossing CH508 and LU648 amoebae (iv) a variety of developing cultures which contained different proportions of fusion cells and zygotes. Some of the developing cultures were obtained by crossing CH508 and LU648 amoebae, while others were obtained from 10-strain matings (2-41%: Table 2.2; Barber, 1998; Bailey et al., 1999; section 2.2.2). The northern blot was probed at high stringency (65°C; section 2.4.4) with the cDNA clones for A18/1020 and actin (pPpA35: Hamelin et al., 1988; Fig. 4.2).

A18/1020 hybridised to a mRNA of approximately 6000nt, consistent with previous findings (section 4.2.1) The A18/1020 mRNA could not be detected in the RNA from amoebae or the developing cell cultures in which 2% and 4% of the cells were committed to development, but was detected in the RNA of the remaining developing cultures (Fig. 4.2a). The 38% and 41% cultures contained the most A18/1020 mRNA (Fig. 4.2a). The A18/1020 mRNA was detected at relatively low levels in the RNA of microplasmodia and macroplasmodia (Fig. 4.2.a). As expected, actin hybridised to a mRNA of 1400nt (Pallotta et al., 1986). All of the samples contained actin transcript that was not significantly degraded.

There are minor differences in the timing of the expression of A18/1020 during apogamic and heterothallic development; the A18/1020 transcript was detected when less apogamic cells were committed to plasmodium development. Similar results were obtained for redA and redB (Bailey et al., 1999). Commitment is operationally defined as the ability of a cell to survive a replating procedure and develop into a plasmodium (sections 1.1.5 and 2.2.2).
In heterothallic development, the cells become committed to plasmodium development at around the time of nuclear fusion, while in apogamic development, the cells become committed to plasmodium development approximately halfway through the extended cell cycle (section 1.1.5). As the relationship between a fusion cell and a uninucleate developing cell has not yet been defined, it is possible that they represent different developmental stages; such a discrepancy could account for the minor differences in gene expression that have been observed. Overall, the expression pattern of A18/1020 is very similar in both apogamic and sexual development; in both cases, A18/1020 is expressed primarily during development. This shows that the changes in gene expression that occur during the APT are similar in both apogamic and heterothallic development, in agreement with the work of Bailey \textit{et al.} (1999).

4.3 SEQUENCING THE A18/1020 cDNA

The cDNA clone was sequenced with the universal 'reverse' and '-40' sequencing primers (Table 2.3), which have binding sites either side of the multiple cloning site within \textit{pBluescript} II KS (Stratagene). The cDNAs were cloned such that the reverse sequencing primer amplified the sense strand from a 5' direction, and the -40 sequencing primer amplified the antisense strand from a 3' direction (section 1.3.3.1). Automated sequencing reactions were performed as described in section 2.4.5. Each primer gave data for approximately 350bp of sequence; a total of 700bp of sequence was determined. As the cDNA clone is 1020bp in length, approximately 320bp of sequence remained to be determined.

To obtain the full sequence of the cDNA, two primers were designed that had binding sites in the already sequenced DNA, A181 and A182 (positions 320-337 and 674-693 respectively: Fig. 4.3; Table 2.3). Purified plasmid was sequenced with these primers (section 2.4.5). Each primer gave data for approximately 300bp of sequence. When the A181 and the A182 sequences were compared, there was a large region of overlap which showed that the complete sequence of the A18/1020 cDNA had been determined.

The clone is 1029bp in length and had 15bp of poly(A)+ tail at its 3' end (Fig. 4.3; EMBL accession number AJ133501). Two of the possible reading frames have stop codons scattered throughout. The other reading frame starts with the third base of the cDNA and has the stop codon, (TAA), located 954bp into the sequence (Fig. 4.3); this must be the ORF. Thus, the cDNA clone contains 953bp of coding sequence, 58bp of 3' untranslated sequence and 15bp of poly(A)+ tail. The consensus polyadenylation signal (aataaa) was identified 18bp upstream of the poly(A)+ tail (position 997-102: Fig. 4.3). The cDNA is not full length because the start codon, ATG, is missing (Fig. 4.3). The deduced protein
sequence contains 317 aa's (Fig. 4.3). The analysis of the sequence is discussed in section 4.5.

4.4 SOUTHERN BLOTTING AND LOW STRINGENCY NORTHERN BLOTTING
Southern blots containing EcoRI, XbaI, AvaI, HindIII, SacI, AvaII and PstI digested CL genomic DNA were probed with the A18/1020 cDNA at high stringency (65°C: sections 2.3.1.3, 2.4.1 and 2.4.4; Fig. 4.4). A single major band was seen in the PstI, EcoRI, XhoI, HindIII and SacI lanes, and two major bands were seen in the AvaI and AvaII lanes because the cDNA contains these restriction sites (Fig. 4.4). Thus, A18/1020 represents a single-copy gene. Other faint bands were visible in every lane of the Southern blots (Fig. 4.4). Although the cDNA contains internal EcoRI and HindIII restriction sites, they are so close to the 5' end of the cDNA clone (<20bp) that the upstream restriction fragments would probably not be detected by Southern blotting. Thus, the Southern blots contained more bands than were predicted from the sequence data.

It is possible that the A18/1020 cDNA spans a region of the gene that has intron(s) containing some of the restriction sites. If any of these restriction enzymes cut within such an intron, two restriction fragments with homology to the cDNA probe would be produced, instead of one, and an extra band would be visualised on the blot. Although some of the extra bands may result from this, it seems unlikely that all of the restriction enzymes cut within such intron(s). It is more likely that the extra bands result from cross-hybridisation of the probe to related sequences because A18/1020 is a member of a gene family. Further evidence for this is discussed in section 4.6.

To see if the A18/1020 cDNA cross-hybridised to related mRNAs, a northern blot like the one described in section 4.2.1 was probed at low stringency (55°C) with the cDNA (data not shown). The cDNA hybridised only to the 6000nt mRNA; no related mRNAs were detected.

4.5 ANALYSIS OF THE A18/1020 PROTEIN
BLAST, FASTA and TFASTA database searches showed that the A18/1020 polypeptide was closely related to the tail domains of type II myosin heavy-chain proteins from many organisms (Fig. 4.5). The A18/1020 protein sequence was most similar to the tail domain of the large myosin heavy-chain protein from Dictyostelium discoideum (Warrick et al., 1986); the proteins were 32% identical, and a further 23% similar over a 315aa region (Fig. 4.5).

Type II myosin molecules are comprised of six polypeptide chains; two identical heavy-chains and two pairs of light-chains (section 1.2.2). The myosin heavy-chain proteins wrap
around each other to give the rod-like myosin tail, and the myosin light-chain proteins attach to the N-terminus of the myosin heavy-chains to form the myosin head (Fig. 4.6a). The myosin heavy-chain proteins adopt unbroken \( \alpha \)-helices that coil around each other to form the coiled-coil structure that is characteristic of the myosin tail (Fig. 4.6a). The two \( \alpha \)-helices will only form a coiled-coil structure if the myosin heavy-chain proteins contain the correct spacing of hydrophobic aa’s. Hydrophobic residues are generally found within positions ‘a’ and ‘d’ of a seven amino acid repeat ‘abcdefg’. The hydrophobic residues of each strand pack tightly to form the hydrophobic core of the myosin tail (Fig. 4.6b). Four heptad repeats are arranged into repeating 28-residue zones that contain regions of alternating charge. Thus, the tail domains of type II myosin heavy-chain proteins have several distinctive structural features. They adopt unbroken \( \alpha \)-helices, they contain heptad repeats with hydrophobic amino acids concentrated at positions ‘a’ and ‘d’, and they have 28-residue zones that contain regions of alternating charge.

An examination of the A18/020 polypeptide showed that it had all of the structural features that are associated with the tail domains of type II myosin heavy-chain proteins. Secondary structure calculations showed that the A18/1020 polypeptide will form an unbroken \( \alpha \)-helix (data not shown). In addition, the protein sequence contains 44 complete heptad repeats, within which hydrophobic residues predominate in positions ‘a’ and ‘d’ (Fig. 4.7a), and has nine complete 28-residue zones that have regions of alternating charge (Fig. 4.7b). The results of the database searches and the structural analyses indicate the partial A18/1020 cDNA encodes part of the tail domain of a type II myosin heavy-chain protein. As northern blotting has shown that A18/1020 is expressed at high levels during development (section 4.1), the cDNA was renamed \textit{mynD} (myosin developmental).

4.6 DISCUSSION

The A18/1020 cDNA clone was fully sequenced. The 1029bp partial cDNA clone contained 953bp of coding sequence, 58bp of 3’ untranslated sequence and 15bp of poly(A)+ tail. This is consistent with studies of other \textit{P. polycephalum} genes which have identified 3’ untranslated regions ranging from 12-195bp (Hamelin et al., 1988; Binette et al., 1990; Bailey et al., 1999). Like many other \textit{P. polycephalum} genes (Nader et al., 1986; Morita, 1998; Bailey et al., 1999), the 3’ untranslated region of A18/1020 contained the consensus polyadenylation signal AATAAA.

The expression of A18/1020 was analysed during apogamic and sexual development by northern blotting. These analyses showed that the cDNA represented a \textit{red} gene that was expressed primarily during both types of development. A18/1020 hybridised to a single
mRNA of approximately 6000nt, indicating the cDNA clone is missing approximately 5kb. The subtracted cDNA library, ML8S, could be rescreened to identify longer A18/1020 cDNA clones. Alternatively, a size fractionated genomic library could be made and screened for an A18/1020 genomic restriction fragment that would contain the missing sequence, as was done for redE (section 3.6). Southern blotting analysis suggested that A18/1020 represented a single-copy gene that is a member of a gene family. To see whether the cDNA hybridised to any related mRNAs, a northern blot containing CL RNA was probed at low stringency (55°C) with the A18/1020 cDNA, but no related mRNAs were detected.

Database searches showed that the A18/1020 cDNA encodes part of the tail domain of a type II myosin heavy-chain protein. Type II myosin has been purified from amoebae and plasmodia (section 1.2.2; Tanigushi et al., 1980; Kohama et al., 1986). An analysis of these molecules showed that the type II myosin heavy-chain proteins of amoebae and plasmodia are different, which suggests amoebae and plasmodia express different type II myosin heavy-chain genes (Kohama et al., 1986). It seems unlikely that mynD encodes the amoebal myosin, because it is not expressed in amoebae. Macroplasmodia contain a lot more myosin than microplasmodia (Ohl and Stockem, 1995). So, presumably the plasmodial myosin gene would be expressed at higher levels in macroplasmodia than in microplasmodia. As mynD is expressed at higher levels in microplasmodia than in macroplasmodia, it seems unlikely that it encodes the major plasmodial myosin. Instead, mynD represents a myosin that is abundant during development.

Laszlo Nyitray and colleagues (Eotvos University) recently identified two cDNA clones (mhciIA and mhciIB) that represent plasmodium-specific type II myosin heavy-chain genes (L. Nyitray, unpublished data). The mhciIA cDNA is full length (6.6kb), while the mhciIB cDNA is not full length (2.6kb) and encodes the myosin head domain. Nyitray and colleagues have evidence which suggest mhciIA encodes the major plasmodial myosin, and mhciIB encodes a minor plasmodial myosin (pers. com.). Sequence comparisons showed that the mynD and the mhciIA cDNAs are different, which confirms that mynD does not represent the major plasmodial myosin. The deduced mynD and mhciIA polypeptides are 78% identical and a further 6% similar. Sequence comparisons showed that mynD and mhciIB cDNAs are different, but since mynD encodes the tail domain of a myosin heavy-chain protein, and mhciIB encodes the head domain of a myosin heavy-chain protein, it is possible that both of the cDNAs represent the same gene. Some basic experiments could be conducted to see if this is the case. A Southern blot could be made which contained genomic DNA digested with a variety of restriction enzymes and the blot could be probed sequentially with the mynD and mhciIB cDNA clones. If both of the cDNAs produced the
same hybridisation pattern, this would indicate that both of the cDNA clones represent the same gene.

Southern blotting suggested that A18/1020 was a member of a gene family. As the MynD and the mhclla deduced polypeptides are 78% identical and a further 6% similar, it seems likely that the \textit{mynD} cDNA cross-hybridised to restriction fragments containing the \textit{mhclla} gene.

Two \textit{Dictyostelium discoideum} cell lines that lack myosin II heavy-chain proteins have been engineered; de Lozanne and Spudich, (1987) achieved gene knockout by homologous recombination with a mutated copy of the myosin gene, while Knecht and Loomis, (1987) achieved gene knockout by antisense RNA gene inactivation. The analysis of these strains showed that although the cells could grow, extend pseudopodia and move by amoeboid locomotion they had several mutant phenotypes; they grew more slowly and only when attached to a substrate, they were defective in cytokinesis and large multinucleate cells were often observed, they were unable to form multicellular aggregates and did not undergo the subsequent morphogenesis, and the rate of pseudopod expansion and the area of the pseudopods was reduced (de Lozanne and Spudich, 1987; Knecht and Loomis, 1987; Wessels \textit{et al}., 1988; Fukui \textit{et al}., 1990 and Traynor \textit{et al}., 1994). Thus, the \textit{Dictyostelium} myosin II heavy-chain protein is required for the correct execution of a range activities including cytokinesis and development.

It seems likely that the \textit{red} genes fall into several categories. Some may be involved in regulating the changes in gene expression that accompany plasmodium development, while others, like \textit{mynD}, may bring about the structural changes that occur during this time. Immunolocalisation experiments, like those proposed for \textit{redE} in section 3.9, would localise the MynD protein in addition to pinpointing the expression of the gene to individual cells. As was proposed for \textit{redE} (section 3.9), the function of \textit{mynD} could be investigated by gene knockout by homologous recombination or antisense experiments or by mis-expressing the gene (see chapter 6). These experiments would shed light on the function of \textit{mynD} and the role of type II myosin in \textit{P. polycephalum}. 
Total RNA was isolated from CL amoebae (A), microplasmodia (Mi), macroplasmodia (Ma) and from a variety of developing cultures (1-56%; Table 2.1) as described in section 2.2.2. 10μg of each total RNA sample, and 3μg of RNA size markers (Promega) were denatured and run on a 1.1% agarose gel as described in section 2.4.2. The RNA was transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.2).

The cDNA clones for (a) A18/1020, (b) proP (Binette et al., 1990) and (c) actin (Hamelin et al., 1988) were radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised sequentially at high stringency (65°C) to the same northern blot. The blot was exposed to x-ray film at -80°C with an intensifying screen. A18/1020 hybridised to a 6000nt mRNA (12 day exposure), proP hybridised to a 520nt mRNA (2 day exposure) and actin hybridised to 1400nt mRNA (1 hour exposure).
Total RNA was isolated from CH508 amoebae (A[CH508]), LU352 amoebae (A[LU648]), microplasmodia (Mi) and macroplasmodia (Ma) formed by crossing CH508 and LU352 amoebae, a variety of developing cell cultures (2-24%) formed by crossing CH508 and LU648 amoebae and two developing cell cultures (40-42%) from ten-strain matings as described in section 2.2.2. Table 2.2 shows the distribution of cell types present in the cultures from which RNA was isolated, and details the strains used in the ten-strain matings. 10μg of each total RNA sample, and 3μg of RNA size markers (Promega) were denatured and run on a 1.1% agarose gel as described in section 2.4.2. The gel was blotted to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.2).

The cDNA clones for (a) A18/1020 and (b) actin (Hamelin et al., 1988) were radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised sequentially at high stringency (65°C) to the same northern blot. The blot was exposed to x-ray film at -80°C with an intensifying screen. A18/1020 hybridised to a 6000nt mRNA (12 day exposure) and actin hybridised to 1400nt mRNA (1 hour exposure).
Figure 4.3: The nucleotide and deduced amino acid sequence of the A18/1020 cDNA clone

The nucleotide sequence of the 1029bp cDNA clone (EMBL accession number AJ133501) was determined by sequencing the A18/1020 cDNA with the universal reverse and -40, A181 and A182 sequencing primers (Table 2.3) as described in section 2.4.5. The sequences were aligned using the Wisconsin Package, Genetics Computer Group (Madison, USA). The positions of the sequencing primers, A181 and A182 are marked on the sequence at bases 320-337 and 674-693 respectively.

The coding region is shown in uppercase letters and the non-coding region is shown in lowercase letters. The open reading frame starts with the third base of the cDNA, and ends with a stop codon at base 954 (TAA), indicated by * in the amino acid sequence. The coding region is followed by 58bp of 3' untranslated sequence and 15bp of poly(A)+ tail. The nucleotide and protein sequences are partial, as the cDNA does not begin with an ATG start codon. The deduced amino acid sequence is shown in standard single-letter code. The putative polyadenylation signal is underlined (ataaa).
AA1  TC GAT AAG CTT GAT ATC GAA TTC CGT TGC TGT CTC CTT CTC GAC GAA

48  TCT TGC GCT ACC AAG ACC CAA TCC GAG AAC CTT GCC GTG AAG CTT GAA

96  GAG GAC ATC AAG AAA CTC CAG GAC GAT TTG GAC AAC GAA ATC AAA AAA

144  AAA GAG CTT ATT GAA AGG ACT CGC AAA TCT CTG GAA CTC CAG CTG GAG

240  GAC AAG CTC CGT CGC CAG GCA GAA AAC GAA TTG GAG GAT TTG CGC GAG

288  CAA GTC GAT GCC CTT GAG GAG ACC ACC CAA GAT CTC GAA TCC ACC AAG

336  ACC CGC CTC GAA GTC GAA TTA GAG GAA GCC AGG AAG TCG AAC GCG CGC

384  GAA AAC GAG GCG GCC GAG GCC GCC GCC GCC GCC GCC GCC GCC GCC GCC GCC

432  CGC GAG TTG GCC GAG TTG CGC GAG AAC GAG GTC ACC GAC ATG GAC GAG GTC

480  R E L R E K S D E E E V I

528  GAC GCC AAG GAG CAG CTG ACC GAC CAG CTG GCC AAG CTT GTC GAC CTA

576  D A K E Q L E T E S K L R A K L

624  ATG AAG GCC GAC GCC GAA AAA CTC CGA TCC AGA ATT TCC TCT TTA GAA

672  AAG ATG GAG GCA GAC TAC AAG AAG ATG CAG GCA CGC TTG ATC GAA GAA

720  K M E A D Y K K M Q A R L I E E

768  GAC ATG AAC GAA GCC CGT TTG GAT GTC ACT TCC ATC AAG GAG GTC ACC

816  GAG AAG ATC CGT GCC GAT CGG GAA AAA GAA AAG CAG GCC ATC GAG GAT

864  E K I R A D A E K E K Q A I E D

912  TTG CAG AAG AAT CTC ACC TCC CTC TCC GGA ATC ACT GCC GGA GCT GCC

961  GCG TGG AAG GCT GGT GCA ACC AGA CCA AGA GAA GAC TAA tttgt

1020  polyadenylation signal  poly(A)+ tail
3μg of genomic DNA from CL microplasmodia was digested with the restriction enzymes shown (section 2.3.1.3) and run on a 0.8% agarose gels with λ HindIII DNA size markers (section 2.3.2). The DNA in the gels was denatured, transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook et al., 1989; section 2.4.1).

The A18/1020 cDNA clone was radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised at high stringency (65°C) to the southern blots. The blots were exposed to x-ray film at -80°C with an intensifying screen for 1-2 weeks. The relative migration of the λ HindIII DNA size markers is shown in bp.
Figure 4.5: Homology between the amino acid sequence of MynD and type II myosin heavy-chain proteins in other organisms (Bailey et al., 1999)

The amino acid sequence of MynD was compared to Dictyostelium discoideum large myosin heavy-chain (Warrick et al., 1986), Oryctolagus cuniculus (rabbit) smooth-muscle myosin (Nagai et al., 1988), Acanthamoeba castellani myosin II (Hammer et al., 1987) and Caenorhabditis elegans myosin II (Dibb et al., 1989). The amino acids are arranged into 28bp repeats. In places gaps (.) were inserted to produce the best fit. The black boxes designate amino acids in MynD that are shared with at least one of the other proteins, while shaded boxes designate amino acids that are similar.
Figure 4.6: The structure of a type II myosin molecule

(a) shows the structure of a type II myosin molecule (redrawn from Alberts et al., 1989).

(b) is a schematic diagram of a transverse section of a myosin tail. Residues 'a' and 'd' within the heptad repeat 'abcdefg' pack tightly to form the hydrophobic core of the myosin tail (redrawn from Stryer, 1988).
Figure: 4.7: Amino acid distribution in MynD

(a) shows the periodicity of hydrophobic residues within MynD. The bars indicate the number of hydrophobic residues found in each position of the heptad repeat 'abcdefg', in 44 complete heptad repeats.

(b) shows the periodicity of charged residues within MynD. The bars indicate the net charge of residues in each position of the heptad repeat 'abcdefg', in nine complete 28 residue regions.

5.1 INTRODUCTION
Most replication studies in *P. polycephalum* have been conducted on genes that are abundantly transcribed in the plasmodium (section 1.6.1). The *red* genes were of particular interest for replication studies because they are expressed at low levels in macroplasmodia (Bailey *et al.*, 1999; section 3.4.1) and genes within this category have not been previously studied. The first aim of this work was to determine the replication timing of the *red* genes in the plasmodium. The transcriptional promoters of early replicating genes that are abundantly transcribed in macroplasmodia often contain DNA replication initiation sites (section 1.6.1). The second aim was to map the replication initiation sites of any early replicating *red* genes, to see if this is true of genes that are not abundantly transcribed in the plasmodium.

The work described in this Chapter was conducted at the ‘Institut de recherches sur le Cancer, CNRS-UPR 1983, 7 rue guy Moquet-BP8, 94801, Villejuif, Cedex, France’, where I worked under the supervision of Dr. Gérard Pierron.

RESULTS

5.2 RESTRICTION MAPPING
A detailed restriction map of the *redB* and *redE* loci was required for the work described in this chapter. Southern blotting was used to create these restriction maps. The mapping of *redE* is described in section 3.5 and the resulting restriction map can be seen in Fig. 3.6. The *redB* locus was mapped in the same way as the *redE* locus. Briefly, Southern blots containing M3CIV digested genomic DNA were prepared and probed with the *redB* cDNA insert (data not shown; sections 2.3.1.3, 2.4.1 and 2.4.4). The resulting restriction map of *redB* can be seen in Fig. 5.1.

A basic restriction map of the *redA*, *proA* and *proP* loci was required for the work described in this chapter. The *proA* and the *proP* mapping data came from Bénard, (1990). Some of this mapping was carried out in the strain, TU291, which is a diploid strain like M3CIV. As no suitable mapping information was available for *redA*, a Southern blot containing EcoRI digested M3CIV genomic DNA was prepared and probed with the *redA* cDNA insert (data not shown; sections 2.4.1 and 2.4.4). The data from this experiment and the *proA* and *proP* mapping data from Bénard, (1990) are shown in Table 5.1. The *redA* cDNA clone contains an EcoRI restriction site approximately 100bp from its 5' end. For this reason, a faint 2.3kb band was visualised on the Southern blot in addition to the major 7.6kb band.
5.3 THE REPLICATION TIMING OF \textit{redA}, \textit{redB} AND \textit{redE} IN THE PLASMODIUM

The replication timing of the \textit{red} genes was analysed using both density shift experiments and quantitative Southern blot analysis.

5.3.1 Density shift experiments

Density shift analysis was carried out as a first step towards discovering when \textit{redA}, \textit{redB} and \textit{redE} are replicated in the plasmodial S-phase. The technique provides a relatively simple way of determining whether a gene is replicated before or after a given point in S-phase. Other analyses can then be carried out to pinpoint the exact time of replication. A brief outline of the technique is given below.

BromodeoxyUridine (BrdU) is an analogue of thymidine that has a higher molecular weight. When BrdU, uridine and fluorodeoxyuridine are supplied to a plasmodium in S-phase, the BrdU is incorporated into the replicating DNA in place of thymidine (section 2.5.2). A plate grown macroplasmodium is placed onto filter paper soaked with BrdU, uridine and fluorodeoxyuridine at the start of S-phase (Fig. 5.2a; section 2.5.2). BrdU is taken into the plasmodium and incorporated into the replicating DNA until the plasmodium is harvested for DNA isolation (section 2.5.3). DNA that was replicated while the plasmodium was in the presence of the drug has BrdU incorporation and is 'BrdU-substituted'. This DNA is heavier than normal DNA because BrdU has a higher molecular weight than thymidine. Due to the semi-conservative nature of DNA replication, the substituted DNA is comprised of one 'light', parental strand of DNA that is non-BrdU substituted and one newly synthesised, 'heavy' strand of DNA that is BrdU-substituted. For this reason BrdU-substituted DNA is referred to as 'heavy light' (HL) DNA. DNA that was not replicated while the plasmodium was in the presence of BrdU is comprised of two 'light' parental strands of DNA that have no BrdU substitution, and is referred to as 'light light' (LL) DNA.

The isolated DNA molecules have replicated regions that are BrdU substituted, interspersed with non-replicated regions that are non-BrdU substituted (Fig. 5.2b). In order to reduce the number of part-substituted DNA molecules, the DNA is digested with a restriction enzyme (section 2.6.1.3). Most of the smaller DNA molecules are now either BrdU-substituted or non-BrdU substituted (Fig. 5.2c). The digested DNA is then centrifuged on a caesium chloride (CsCl) gradient to separate the normal, LL DNA from the density shifted, HL DNA (section 2.7.2). After centrifugation, the LL and the HL DNA fractions can be easily distinguished (Fig. 5.2d; Fig. 5.3) and they are isolated separately from the centrifuge tube with a syringe (Fig. 5.2e; section 2.7.2). The two DNA fractions
are never completely pure because this method of separation is relatively crude. Rather, there is an enrichment of LL and HL DNA in the respective samples.

Approximately 5μg of the LL and HL DNA is run on an agarose gel with λ HindIII and φX174 HaeIII DNA size markers (section 2.7.2). After running, the gel is photographed under UV illumination and transferred to membrane. The blot is then probed for the gene of interest (section 2.7.4). If the probe hybridises to DNA fragments in the unreplicated, LL DNA fraction, the gene was not replicated by the time the plasmodium was harvested. In contrast, if the probe hybridises to DNA in the replicated, HL DNA fraction, the gene was replicated during the time the plasmodium was in the presence of BrdU. In order to check the quality and separation of the LL and HL DNA fractions, it is usual to hybridise blots with at least two probes; one for the gene of interest and one for a gene whose replication timing is already known. If the known probe hybridises predominantly to DNA in the expected fraction, the separation of LL and HL DNA is reasonable and any result obtained for the gene of interest will be valid.

5.3.1.1 The replication timing of \textit{redA} and \textit{proA}
A blot was made which contained M3CIV EcoRI digested G2-phase DNA, and LL and HL DNA from a plasmodium that had been treated with BrdU for the first 40 min of S-phase (section 5.3.1). The blot was probed for \textit{redA} and \textit{proA} (Fig. 5.4b; see Table 5.1 for the expected \textit{redA} and \textit{proA} band sizes). The amoeba-specific profilin gene, \textit{proA}, was used as a control because it is known to replicate 40-60 min from the onset S-phase in the plasmodium (Bénard and Pierron, 1992).

As expected, \textit{proA} hybridised to two bands because of an allelic polymorphism (Fig. 5.4b; Table 5.1). The 2.3kb \textit{redA} restriction fragment was not detected on this blot (Fig. 5.4b; section 5.2). When restriction fragments containing \textit{redA} and \textit{proA} are equally abundant, both of the probes produce bands of roughly the same intensity (G2: Fig. 5.4b). This shows that the probes are of equal strength and that any intensity differences between the \textit{redA} and \textit{proA} bands in the LL and HL DNA will reflect a shift in the number of restriction fragments containing each gene. Both \textit{redA} and \textit{proA} have hybridised to DNA in the LL and the HL DNA fractions, which indicates the fractions are not particularly pure (Fig. 5.4b). In the LL DNA sample the \textit{redA} band is stronger than either of the \textit{proA} bands and the reverse of this is seen in the HL DNA sample, where the \textit{proA} bands are more intense than the \textit{redA} band (Fig. 5.4b). This shows that most of the restriction fragments containing \textit{proA} are in the replicated, HL DNA fraction, while most of the restriction fragments containing \textit{redA} are in the unreplicated, LL DNA fraction. Thus, \textit{proA} has been replicated and \textit{redA} has not.
The data suggest that \textit{proA} is replicated during the first 40 min of S-phase, yet it is known that \textit{proA} is replicated between 40-60 min from the onset of S-phase (Bénard and Pierron, 1992). It seems likely that the DNA fractions were isolated from a plasmodium that was treated with BrdU for longer than 40 min. However, this does not affect the analysis of \textit{redA}, which is clearly replicated later than \textit{proA}. These results suggest that \textit{redA} is a late replicating gene that is not replicated during the first hour of S-phase in the plasmodium.

5.3.1.2 The replication timing of \textit{redB} and \textit{proP}

A blot was made which contained M3CIV \textit{HindIII} digested G2-phase DNA, and LL and HL DNA from a plasmodium that had been treated with BrdU for the first 40 min of S-phase (section 5.3.1). The blot was probed for \textit{redB} and \textit{proP} (Fig. 5.4d; see Fig. 5.1 and Table 5.1 for the expected \textit{redB} and \textit{proP} band sizes). The plasmodium-specific profilin gene, \textit{proP}, was used as a control because it is known to replicate at the onset S-phase in the plasmodium (Bénard and Pierron, 1992).

When restriction fragments containing \textit{redB} and \textit{proP} are equally abundant, both of the probes produce bands of roughly the same intensity (G2: Fig. 5.4d). Which shows that the probes are of equal strength and that any intensity differences between \textit{redB} and \textit{proP} bands in the LL and HL DNA will reflect a shift in the number of restriction fragments containing each gene. Both \textit{proP} and \textit{redB} produced much stronger bands in the replicated, HL DNA than in the unreplicated, LL DNA (Fig. 5.4d) which suggests that both genes are replicated during the first 40 min of S-phase. The data seems to be reliable because \textit{proP} has given the expected result. Thus, \textit{redB} is an early-replicating gene that is replicated during the first 40 min of S-phase in the plasmodium.

5.3.1.3 The replication timing of \textit{redE} and \textit{proP}

A blot was made which contained TU291 \textit{EcoRI} digested LL and HL DNA from a plasmodium that had been treated with BrdU for the first 45 min of S-phase (section 5.3.1). The blot was probed for \textit{redE} and \textit{proP} (Fig. 5.5b; see Fig. 5.1 and Table 5.1 for the expected \textit{redE} and \textit{proP} band sizes).

As expected, \textit{redE} hybridised to two bands because of an internal \textit{EcoRI} restriction site (Fig. 3.6; Fig. 5.5b). Both \textit{proP} and \textit{redE} hybridised more strongly to the replicated, HL DNA, than to the unreplicated, LL DNA which suggests both genes are replicated during the first 45 min of S-phase. The data are reliable because \textit{proP} has given the expected result. Thus, \textit{redE} is replicated during the first 40 min of S-phase in the plasmodium.
5.3.1.4 A summary: the replication timing of redA, redB, redE and proP

The aim of this experiment was to summarise the redA, redB and redE density shift data in a single experiment. A blot was made which contained M3CIV EcoRI digested LL and HL DNA from a plasmodium that had been treated with BrdU for the first 45 min of S-phase (section 5.3.1). The blot was hybridised, in a single hybridisation, to a mixture of four probes (redA, redB, redE and proP; Fig. 5.5d; see Figures 3.6 and 5.1 and Table 5.1 for the expected band sizes).

All of the probes hybridised to DNA in the replicated, HL DNA fraction, except for redA, which hybridised to DNA in the unreplicated, LL DNA fraction (Fig. 5.5d). It is likely that the 2.3kb band in the LL DNA fraction is the result of redA hybridisation (section 5.2; Table 5.1), although it is possible that some of the signal results from redB hybridisation because restriction fragments containing redB are not completely density shifted. The data confirms that redB, redE and proP are replicated during the first 45 min of S-phase in the plasmodium, and that redA is replicated at some later point.

5.3.2 Quantitative Southern blot analysis

Although density shift analysis established approximately when in S-phase the red genes were replicated, their replication timing was also examined using quantitative Southern blot analysis. This type of analysis has two advantages over density shift analysis; it pinpoints the replication timing of a gene more accurately and the plasmodia are not treated with drugs that could potentially disrupt the temporal order of replication.

Prior to the onset of S-phase each nucleus in a plasmodium has two copies of each single-copy gene and after DNA replication, when S-phase is complete, four copies of each gene. Any two restriction fragments containing single-copy genes will be present in equal numbers, in a 1:1 ratio in either the unreplicated or the replicated genome. If one DNA fragment replicates before the other, it will be twice as abundant as the unreplicated fragment and the ratio will shift to 2:1 for a period of time during S-phase. This two-fold increase in copy number can be detected by measuring the abundance of one DNA fragment relative to the other, using quantitative Southern blot analysis (Pierron et al., 1984 and 1989). When the second DNA fragment is replicated, the fragments will again be equal in number and the abundance ratio will return to 1:1.

5.3.2.1 The replication timing of redA, redB and proP

DNA was extracted from a plasmodium in G2-phase of the cell cycle and from six different plasmodia 30 min, 60 min, 90 min, 120 min, 180 min and 300 min after the onset of S-phase (sections 2.5.1 and 2.5.3). The DNA was digested with EcoRI (section 2.6.1.2), electrophoresed (Fig. 5.6a; section 2.6.2) and transferred to membrane (section 2.7.1). The
blot was then probed for redA, redB and proP (Fig. 5.6b; section 2.7.4; see Fig. 5.1 and Table 5.1 for the expected band sizes; section 2.7.4). Three bands, which correspond to restriction fragments containing each gene, can be seen in every lane. The plasmodium-specific profilin gene, proP, was used as a control because it is known to replicate at the onset of S-phase (Bénard and Pierron, 1992). After probing the blot was exposed to a special screen for 4-8 hours which was scanned with a phosphorimager (section 2.7.4).

There were several steps in the analysis of the blot:

(i) The intensity of each band on the Southern blot was quantitatively measured from the scanned image using the application 'Image Quant'. Each band was given a 'volume report' in arbitrary units (data not shown). There was a positive correlation between the volume report and band intensity; the more intense the band, the higher the volume report.

(ii) The G2 volume reports were called 1.0 because restriction fragments containing redA, redB and proP are equally abundant in this sample. The volume reports in all other lanes were then normalised to the volume reports in this lane, e.g. the redA G2 volume report is 58389 and the redA 30' volume report is 47266. The normalised 30' volume report is therefore 0.81 (58389/47266). The remaining redA volume reports were normalised to the G2 volume report in this way and the procedure was repeated for proP and redB.

(iii) The normalised volume reports were then compared to give band ratios, e.g. the normalised proP 30' volume report is 1.6 and the normalised redA 30' volume report is 0.81. The 30' proP/redA band ratio is therefore 1.98 (1.6/0.81). The remaining proP/redA band ratios were calculated in this way and the procedure was repeated for proP/redB and redB/redA. These band ratios were plotted on graphs (see below). Band intensities were only compared within a gel lane and never between gel lanes to ensure that when the abundance of one restriction fragment was compared to the abundance of another, both of the restriction fragments originated from the same DNA sample.

It was necessary to normalise the volume reports in each lane to the volume reports in the G2 sample, because the probes were not equally sensitive. This had to be taken into account before the band intensities of different genes could be compared. The band ratios ranged from 1.0-2.0. A ratio of 1.0 showed that restriction fragments containing both genes were equally abundant in that sample, while a ratio of 2.0 showed that restriction fragments containing the first gene were twice as abundant as restriction fragments containing the second gene.
The proP/redA band ratios are shown in Figure 5.7a. The 180 and 300 min DNA samples have band ratios which are close to 1.0, which shows that restriction fragments containing proP and redA are equally abundant in these samples. This is the expected result because these samples should consist of fully replicated DNA. In the 30 min sample the band ratio is close to 2.0, which shows that restriction fragments containing proP are almost twice as abundant as restriction fragments containing redA. This suggests that proP has been replicated during the first 30 min of S-phase and that redA has not. In the 60 min sample the band ratio is 1.7, which shows that restriction fragments containing proP are no longer twice as abundant as restriction fragments containing redA. This suggests that some of the restriction fragments containing redA have been replicated. In the 90 min sample the band ratio is 1.2, which shows that restriction fragments containing redA have been replicated. The band ratios are close to 1.0 in the remaining samples. This data confirms that proP is replicated during the first 30 min of S-phase (Bénard and Pierron, 1992) and shows that most of the restriction fragments containing redA are replicated 60-90 min after the onset of S-phase.

The proP/redB band ratios are shown in Figure 5.7b. The band ratios are always close to 1.0, which shows that restriction fragments containing redB and proP are equally abundant in every DNA sample. This suggests proP and redB are replicated in the same temporal window. As proP is known to replicate during the first 30 min of S-phase (Bénard and Pierron, 1992), redB must also replicate during this time.

The redB/redA band ratios can be seen in Figure 5.6c. As expected, the samples containing fully replicated DNA (180 and 300 min) have band ratios which are close to 1.0. In the 30 min sample the band ratio is 1.7, which shows that restriction fragments containing redB are approximately 1.7 times more abundant than restriction fragments containing redA. This suggests that redB has been replicated during the first 30 min of S-phase and that redA has not. In the 60 min sample the band ratio is 1.6, which shows that some of the restriction fragments containing redA have been replicated. In the 90 min sample the band ratio is 1.2 which shows that restriction fragments containing redB and redA are almost equal in abundance. This suggests that most of the restriction fragments containing redA have been replicated. The band ratio is not significantly different from 1.0 in any subsequent sample.

Previous studies have shown that a maximum band ratio variance of ± 10% in identical lanes can be expected in this type of analysis (Cunningham and Dove, 1993). By this rule the 30 min redB/redA band ratio of 1.73 is significantly different from 2.0. Usually this result would suggest that 30 min into S-phase either (i) redB replication is not complete, or
(ii) some of the restriction fragments containing redA have been replicated. But neither of these scenarios seems likely because (i) the \( \text{proP/redB} \) band ratios suggest that the replication of redB is complete 30 min into S-phase and (ii) the \( \text{proP/redA} \) band ratios suggest that redA replication does not start during the first 30 min of S-phase. The redA cDNA hybridises weakly to a 2.3kb EcoRI restriction fragment on Southern blots (data not shown; section 5.2; Table 5.1). If, at 30 min, redB replication is complete and redA replication has not started, the 2.3kb band could be the result of hybridisation to replicated copies of redB and unreplicated copies of redA. In this scenario the 30 min \( \text{redA/redB} \) band abundance ratio could never be 2.0, because a fraction of the 2.3kb band represents unreplicated redA restriction fragments. This seems the most likely explanation. In summary, the \( \text{redB/redA} \) band ratios show that most of the restriction fragments containing redA are replicated 60-90 min from the onset of S-phase. In contrast, redB is replicated during the first 30 min of S-phase.

The data suggest that redA is replicated over a broad temporal window. The \( \text{proP/redA} \) and the \( \text{redB/redA} \) band ratios show that redA replication starts 30-60 min from the onset of S-phase and is complete by 180 min, at the end of S-phase. The biggest drop, in both the \( \text{proP/redA} \) and the \( \text{redB/redA} \) band ratios, occurs between 60-90 min from the onset of S-phase, which shows that most of the restriction fragments containing redA are replicated during this time. The data is in agreement with the density shift analysis which showed that redA is not replicated during the first 45 min of S-phase. The fact that redA is replicated in mid S-phase shows that the origin(s) that result in the replication of redA are either far from the gene or late firing. The gene was not studied further. It is clear from both the \( \text{proP/redB} \) and the \( \text{redB/redA} \) band ratios that redB is replicated during the first 30 min of S-phase. This is in agreement with the density shift experiments which showed that redB is replicated during the first 40 min of S-phase.

### 5.4 2D GEL ELECTROPHORESIS AND SOUTHERN HYBRIDISATION

The next step was to pinpoint the exact time of replication of redB and redE, and to map any replication initiation sites that may lie close to the genes by 2D gel electrophoresis and Southern hybridisation. A brief account of the technique is given below.

The replication of a single-copy locus can be visualised using 2D gel electrophoresis and Southern hybridisation. Many restriction fragments are generated when genomic DNA from a plasmodium in S-phase is digested with a restriction enzyme. Most of the DNA molecules will be linear because they were not replicating at the time the plasmodium was harvested. A small proportion of the restriction fragments will be more structurally
complex because they were replicating at the time the plasmodium was harvested; they will contain replication bubble(s) and/or replication fork(s). These DNA molecules are sometimes referred to as ‘replication intermediates’.

DNA is extracted from a plasmodium in S-phase (section 2.5.3), imbedded in an agarose plug (section 2.5.3.1) and digested with a restriction enzyme (Fig. 5.8a; section 2.6.1.4). The DNA in the agarose plug is then run on a 0.4% agarose gel (Fig. 5.8b; section 2.7.3). In this dimension, the replication intermediates migrate more slowly than the corresponding linear restriction fragments because of their size. After running, the gel lane containing the size fractionated genomic DNA is excised from the gel (Fig. 5.8c) and placed horizontally across the top of a gel bed, in the position the lanes usually occupy. Melted 1.1% agarose is poured around the 0.4% gel slice to form a new gel, which is run under conditions of high voltage (Fig. 5.8d; section 2.7.3). In this dimension, the replication intermediates migrate more slowly than the corresponding linear restriction fragments because of their size and branched structure. The linear DNA molecules form a continuous diagonal curve on a 2D gel; the small DNA molecules run far into the gel (Fig. 5.8f) and the larger DNA molecules run less far into the gel (Fig. 5.8e). This curve can be visualised when the gel is stained with ethidium bromide and examined under UV illumination. The replication intermediates are located above the line of linear DNA because they migrate more slowly than the corresponding linear DNA fragments in both the first and second dimensions, but they cannot be visualised under UV illumination because their concentration is too low. There are three basic types of replication intermediates. A restriction fragment can contain (i) a replication bubble (ii) a replication fork or (iii) two converging replication forks. Each has a characteristic pattern of migration on a 2D gel.

Single-copy restriction fragments are located on 2D gels by Southern hybridisation (sections 2.7.3 and 2.7.4). If the restriction fragment was not replicating at the time the plasmodium was harvested, a ‘1x spot’ will be visualised on the blot. This spot of signal will be located on the line of linear DNA. Its exact positioning will depend on the size of the restriction fragment (Fig. 5.8). If the restriction fragment was replicating at the time the plasmodium was harvested, a more complex pattern of hybridisation will be visualised on the blot. Schematic diagrams of the hybridisation patterns that are produced by four types of replication intermediates are shown in Fig. 5.9. Each depicts the replication of a hypothetical 3kb restriction fragment.

Firstly, a restriction fragment that is replicated by a single replication fork is visualised as a ‘simple Y’ arc (Fig. 5.9a). The linear DNA is visualised at the 1x spot (marked o, 3kb: Fig. 5.9a). As the replication fork enters the restriction fragment it moves away from the
line of linear DNA as its size and structure changes. The restriction fragment migrates more slowly and becomes larger as the replication fork progresses. When the replication fork is approximately halfway through the restriction fragment it migrates the most slowly in the second dimension. As the replication fork passes through the fragment, it runs more and more like linear DNA. When replication is almost complete, the restriction fragment is almost 6kb and runs at the line of linear DNA. The point at which the restriction fragment touches this line is called the ‘2x spot’ (marked o, 6kb: Fig. 5.9a). Two 3kb restriction fragments are generated when the replication fork exits the fragment at the 2x spot.

Secondly, a restriction fragment that is replicated by a centrally located replication bubble is visualised as a ‘bubble’ arc (Fig. 5.9b). As the replication bubble forms and grows the restriction fragment moves away from the line of linear DNA as its size and structure changes. The restriction fragment migrates more slowly and becomes larger as the replication bubble grows. When the replication bubble is as big as it can be without having seen the exit of one replication fork, it migrates the most slowly in the second dimension. The restriction fragment does not return to the line of linear DNA until two 3kb restriction fragments are generated when the replication forks exit either end of the restriction fragment at the apex of the bubble arc.

Thirdly, a restriction fragment that is replicated by an asymmetrically located replication bubble is visualised as a ‘bubble to Y’ transition (Fig. 5.9c). As the replication bubble forms and grows the restriction fragment moves away from the line of linear DNA as its size and structure changes. The restriction fragment migrates more slowly and becomes larger as the replication bubble grows. When the replication bubble is as large as it can be without having seen the exit of one replication fork, it migrates the most slowly in the second dimension. When one of the replication forks exits the restriction fragment the bubble arc is converted to a simple Y arc. As the remaining replication fork passes through the fragment, it runs more and more like linear DNA and eventually reaches the ‘2x spot’.

Finally, a restriction fragment that is replicated by two converging replication forks is visualised as a ‘double Y’ (Fig. 5.9d). As the replication forks enter the restriction fragment from either end, it moves away from the line of linear DNA as its size and structure changes. The restriction fragment migrates more slowly and becomes larger as the replication forks advance. The restriction fragment migrates the most slowly in the second dimension just before the replication forks meet and the molecule is X shaped (Fig. 5.9d). The restriction fragment does not return to the line of linear DNA until two 3kb restriction fragments are generated when the replication forks meet in the middle of the fragment, at the apex of the double Y.
In some cases, two replication forks clash off-centre in a restriction fragment because they have entered the restriction fragment at different times. If one replication fork enters a restriction fragment before the other, the start of a simple Y arc will be visualised. When the second replication fork enters the restriction fragment, the simple Y arc is converted to a double Y. As the second replication fork progresses through the restriction fragment, it moves away from the simple Y arc because it migrates more slowly in the second dimension. The replication of restriction fragments that contain two replication forks ends somewhere along the x spike (marked x: Fig. 5.9d). The replication of a restriction fragment which contains replication forks that clash slightly off-centre, will end near the top of the x spike. While, the replication of a restriction fragment which contains replication forks that clash very off-centre, will end nearer the bottom of the x spike. The apex of a double Y is sometimes referred to as a ‘termination signal’ because it indicates the clashing of two replication forks.

Arcs are visualised on blots of 2D gels because there is a slight asynchrony in the plasmodial S-phase (approximately 5 min: Bénard and Pierron, 1990; section 1.6.1). If the restriction fragment of interest was replicated at exactly the same time in every nucleus, a ‘spot’ of signal would be seen somewhere on the arc. A restriction fragment which is replicated by a replication bubble is said to be ‘actively replicated’, while a restriction fragment that is replicated by replication fork(s) is said to be ‘passively replicated’. A restriction fragment that contains a very asymmetrically located replication origin will be visualised as a simple Y arc instead of a bubble to Y transition. This is because the start of the ascending arms of simple Y and bubble arcs are hard to distinguish (Fig. 5.9a and b). While the replication bubble is still very small, the fragment sees the exit of one replication fork which converts the bubble arc to a simple Y arc. The replication of overlapping restriction fragments is sometimes analysed to see if the restriction fragment that generates a simple Y arc contains such an origin.

5.4.1 Analysis of redB replication using 2D gel electrophoresis
Three main DNA samples were used to analyse redB replication by 2D gel electrophoresis. They were called +4', +10' and +15' and were isolated from plasmodia 4 min, 10 min and 15 min after the onset of S-phase. When a Southern blot containing PstI digested genomic DNA is probed with the redB cDNA, two bands of 8.5kb and 6.4kb are produced (data not shown) because of a PstI site within the gene. The 8.5kb PstI restriction fragment stretches upstream of redB, while the 6.4kb PstI restriction fragment stretches downstream of redB (Fig. 5.1). Throughout this chapter, these restriction fragments are referred to as the ‘upstream PstI fragment’ and the ‘downstream PstI fragment’ respectively.
Some of the +4', +10' and +15' DNA was digested with *PstI* and run on three 2D gels. The DNA was transferred to membrane, and the blots were probed for *redB* (Fig. 5.10). When looking at the *PstI* 2D gel series (Fig. 5.10) it is clear that the locus is replicating just 4 min after the onset of S-phase. This shows that *redB* is replicated at the onset of S-phase in the plasmodium in agreement with the density shift and quantitative Southern blotting analysis. The 2D gel series shows that the upstream *PstI* restriction fragment is replicated slightly before the downstream *PstI* restriction fragment. At +4' the upstream *PstI* fragment is replicating as the ascending arm of a simple Y arc can be seen (Fig. 5.10). But the downstream *PstI* fragment is not replicating as only the '1x spot' can be visualised (Fig. 5.10). At +10' the replication of the upstream fragment has progressed and apex of the simple Y arc is visible (Fig. 5.10). While the replication of the downstream *PstI* fragment has started as the ascending arm of a simple Y arc can be seen (Fig. 5.10). At +15' the replication of the upstream *PstI* fragment is nearly complete as the descending arm of the Y arc can be visualised, down to the '2x spot' (Fig. 5.10). Most of the downstream *PstI* fragments are about half replicated at this time because the strongest hybridisation is to the apex and the start of the descending arm of the simple Y arc (Fig. 5.10).

The replication of the upstream and the downstream *PstI* fragments has generated two clear simple Y arcs on the +15' blot (Fig. 5.10). This shows that both of the restriction fragments are passively replicated and that neither of them contain a centrally located active replication origin. The fact that *redB* is replicated at the onset of S-phase shows, however, that the gene is close to active replication origin(s) that fire at the onset of S-phase. The *PstI* gel series shows that the upstream and the downstream *PstI* restriction fragments are not replicated by the same replication fork. If they were, a replication fork would have to pass through the first restriction fragment before it could enter the second; the replication of one restriction fragment would have to end before the replication of the second could begin. As the two *PstI* fragments replicate simultaneously (to within 5 min) this cannot be the case. Instead, the restriction fragments are replicated by two different replication forks.

If the *PstI* fragments are replicated by two different replication forks there are two possible scenarios. Either (a) an origin located within the gene region generates two replication forks that diverge from *redB* (a: Fig. 5.10). We would not expect to see bubble arcs on either of the *PstI* fragments if this were the case, because the origin would lie very close to one end of both of the restriction fragments. Or (b) two origins that lie either side of *redB* generate replication forks that converge towards *redB* (b: Fig. 5.10). We would expect to see a termination signal near the gene if this were the case, because when both origins fired
on the same chromatid, they would generate replication forks that would clash in the gene region.

A test was designed to see which of the scenarios was correct. If replication forks diverge from \textit{redB}, the replication forks seen at +4' in the upstream \textit{PstI} fragment will be in the 3' region of the fragment that overlaps the gene (Fig. 5.10). On the contrary, if replication forks converge towards \textit{redB}, the replication forks in the upstream \textit{PstI} fragment at +4' will be in the 5' end of the restriction fragment and will not overlap with the gene (Fig. 5.10). To see where replication forks are located at this time, the replication of the \textit{HindIII} and \textit{KpnI} restriction fragment was examined at +4'. This digest was chosen because it generates a \textit{redB} restriction fragment which spans the gene region. The following predictions were made; (a) if replication forks diverge from \textit{redB} the \textit{HindIII KpnI} restriction fragment will be replicating at +4' (prediction a: Fig. 5.10) and (b) if replication forks converge towards \textit{redB} the \textit{HindIII KpnI} restriction fragment will not be replicating at +4' (prediction b: Fig. 5.10). There was a '1x spot' on the blot (Fig. 5.10). As the \textit{HindIII KpnI} restriction fragment is not replicating at +4', an active origin cannot be located within \textit{redB}. Instead, two replication forks must converge towards \textit{redB}. For this to occur, two active origins must lie either side of \textit{redB}. These origins must both fire at the onset of S-phase for \textit{redB} to be replicated at this time.

Some further experiments were conducted to confirm this result. As previously mentioned, if the two origins that lie either side of \textit{redB} fire on the same chromatid, we would expect replication forks to clash in the gene region. To see if this was occurring, the replication kinetics of the \textit{HincII} and \textit{XhoI} restriction fragments was examined (Fig. 5.11). At +4' the \textit{HincII} restriction fragment is not replicating as all that can be seen on the blot is the 1x spot (Fig. 5.11). In contrast, the \textit{XhoI} fragment is starting to replicate as the ascending arm of a simple Y arc can be seen (Fig. 5.11). The fact that the \textit{XhoI} fragment is replicating before the \textit{HincII} fragment shows that the replication forks have entered the \textit{XhoI} fragment from a 3' direction. If the forks had entered from a 5' direction, the \textit{HincII} fragment would be replicating before the \textit{XhoI} fragment; and this is not the case. At +10' replication of the \textit{HincII} fragment has begun and the ascending arm of a simple Y arc can be seen (Fig. 5.11). Meanwhile, replication of the \textit{XhoI} fragment has progressed and the apex of the Y arc can be visualised (Fig. 5.11). So at +10' replication forks have just entered the \textit{HincII} fragment, while the \textit{XhoI} fragment is approximately half replicated. At +15' a clear termination signal can be seen on both the \textit{HincII} and the \textit{XhoI} blots (Fig. 5.11). This shows that replication forks have entered the \textit{HincII} and the \textit{XhoI} restriction fragments from both ends have clashed. This data confirms that active replication origins lie either side of \textit{redB}. When both origins fire on the same chromatid, they generate replication forks.
that converge towards redB, and clash somewhere between the 5' Xhol and the 3' HincII restriction sites.

The rate of fork movement through the 8.5kb upstream PstI fragment was determined by analysing the PstI 2D gel series (Fig. 5.10). On the +4' blot the ascending arm of a Y arc can be seen, not quite to the apex (Fig. 5.10A). This suggests that replication forks are approximately 1-2kb through the 8.5kb restriction fragment. At +10' the apex of the Y arc can be seen, which suggests that replication forks are approximately halfway, or 4.5kb, through the restriction fragment (Fig. 5.10A). At +15' the descending arm of the Y arc can be seen, down to the '2x spot' (Fig. 5.10A). This suggests that replication forks are approximately 6.5-8kb through the 8.5kb restriction fragment. So the replication forks in the upstream PstI fragment travel from 1.5kb at +4' to 4.5kb at +10' to 7kb at +15', covering approximately 5.5kb in 11 min, or 0.5kb/min. This rate of fork movement is similar to that observed in previous studies. The kinetics of elongation P. polycephalum replicons has previously been measured at approximately 1.2kb/min, or 0.6kb/min/fork (Funderaud and Haugli, 1978a and 1979).

As previously discussed, replication forks are clashing roughly in the middle of the HincII restriction fragment at +15'. Assuming that both origins fire at the onset of S-phase and taking into account the rate of fork movement of 0.5kb/min, it can be estimated that the upstream and the downstream origins are located 7.5kb from the middle of the HincII restriction fragment. This places the upstream origin approximately 1kb downstream of the 5' PstI site, and the downstream origin approximately 1kb downstream of the 3' Xhol restriction site (Fig. 5.11).

Collectively the data indicate that there are two replication origins located close to redB; one upstream and one downstream. When both origins fire on the same chromatid, they produce replication forks that converge towards the gene and clash in the region directly 3' of redB. Both of the origins fire at the onset of S-phase in the plasmodium. The rate of replication fork movement was estimated to be 0.5kb/min from the PstI 2D gel series. This rate of fork movement places the two origins approximately 6kb upstream and 8kb downstream of the gene.

5.4.2 Analysis of redE replication using 2D gel electrophoresis
DNA was extracted from a plasmodium that was 10 min into S-phase. This DNA sample was called +10' and was used to make several 2D gels for redE. The +10' DNA was digested with BglII, KpnI and EcoRV and the digested DNAs were run on 2D gels. The resulting blots were probed for redE (Fig. 5.12). All three restriction enzymes generate restriction fragments that span the region downstream of redE (Fig. 3.7). Due to an allelic
polymorphism in the diploid strain, M3CIV, an EcoRV digest generates two restriction fragments containing redE, sized 9kb and 15kb. Only the replication of the smallest EcoRV allele will be discussed in this section. A simple Y arc can be seen on the BglII blot (Fig. 5.12), while apparent bubble to Y transitions can be seen on the KpnI and EcoRV blots (Fig. 5.12). The occurrence of bubble arcs on the KpnI and EcoRV blots shows that both of the restriction fragments are actively replicated. This shows that a replication origin is located between the 5’ EcoRV and the 3’ KpnI restriction sites (Fig. 5.12). For the KpnI and the EcoRV restriction fragments to be replicating at +10’, the origin must fire at the onset of S-phase in the plasmodium.

Looking at the +10’ 2D gel series gives clues to the positioning of the replication origin. On the +10’ BglII blot only a Y arc can be seen (Fig. 5.12). The absence of a bubble arc indicates the origin is either outside, or close to one edge of this restriction fragment. There is a bubble to Y transition on the +10’ KpnI blot (Fig. 5.12). The strongest hybridisation is to the apex and the start of the descending arm of the simple Y arc. This indicates that most of the KpnI fragments have just seen the exit of one replication fork and are half replicated. There is also a bubble to Y transition on the +10’ EcoRV blot. The strongest hybridisation is to the apex of the bubble and simple Y arcs (Fig. 5.12). This indicates that many EcoRV restriction fragments still contain a replication bubble, although some have just seen the exit of one replication fork. These observations suggest the origin is more centrally located on the EcoRV restriction fragment than it is on the KpnI restriction fragment. For this to be true, the origin must be located towards the 3’ end of the KpnI restriction fragment near the 3’ BglII site (Fig. 5.12). The origin is activated at the onset of S-phase giving rise to a replication bubble. A replication fork exits, firstly the 3’ KpnI restriction site, and secondly the 3’ EcoRV restriction site.

To see what was occurring upstream of the gene, two 2D gels were made with HindIII and EcoRI digested DNA from a plasmodium that was 15 min into S-phase. The resulting blots were probed for redE (Fig. 5.12). Simple Y arcs can be seen on the blots and the complete replication of both of the restriction fragments can be seen. This shows the HindIII and the EcoRI restriction fragments are passively replicated and that neither contain a centrally located replication origin. Interestingly termination signals can be seen on both of the blots. So at +15’, replication forks that have entered both ends of the HindIII and the EcoRI restriction fragments have clashed. For this to occur there must be an active origin located upstream of redE, in addition to the 3’ replication origin. For replication forks to clash at +15’, this 5’ replication origin must be close to the upstream EcoRI site, approximately 8kb upstream of the gene and, like the 3’ replication origin, fire at the onset of S-phase.
The blots depicting the replication of the *KpnI* and *EcoRV* restriction fragments (Fig. 5.12) support the idea that a second replication origin is located close by. The replication of these two fragments generates a complete simple Y arc, in addition to a bubble arc. This is termed a ‘composite signal’ and is different from a bubble to Y transition because the ascending arm of the Y arc is visible. Composite signals are indicative of a restriction fragment being both actively and passively replicated. The existence of composite signals on the *KpnI* and *EcoRV* fragments suggests the 3’ replication origin is ‘weak’, because if it fired 100% of the time giving rise to a clear bubble to Y transition, the ascending arm of the Y arc would never be seen. Some DNA samples give composite signals instead of clear bubble to Y transitions for well-defined origins. It is believed that replication bubbles that are sometimes randomly broken during the experimental procedure can run along the simple Y arc axis generating confusing data (G. Pierron, pers. comm.; Linskens and Huberman, 1990). To determine whether the composite signals were an artefact of the +10’ DNA sample, some control experiments were carried out.

The first experiment was designed to check that the +10’ DNA sample does not replace bubble to Y transitions with composite signals. The experiments used the plasmodium-specific profilin gene, *proP*. The replication of this locus produces a bubble to Y transition on 2D gels because the gene is tightly linked to a replication origin that always fires at the onset of S-phase (Bénard and Pierron, 1992). The +10’ DNA was digested with a restriction enzyme that gave *proP* on a restriction fragment containing its origin, run on a 2D gel and the resulting blot was probed for *proP*. A clear bubble to Y transition can be seen on the blot (Fig. 5.13b) and the signal is not composite because the ascending arm of a Y arc cannot be seen. This data suggests that the composite signals are specific to the *redE* locus, because the +10’ DNA sample does not replace bubble to Y transitions with composite signals.

The second experiment was designed to further confirm that the composite signals were specific to the replication of *redE*. A new DNA sample, +5’, was digested with *KpnI*, run on a 2D gel and the resulting blot was probed for *redE* (Fig. 5.13c). There is again a bubble arc and a complete simple Y arc on the blot. The signal is definitely composite because the ascending arm of the simple Y arc can be seen. The +5’ DNA was then digested with a restriction enzyme that gave *proP* on a restriction fragment containing its origin, run on a 2D gel and the resulting blot was probed for *proP* (Fig. 5.13d). A bubble to Y transition can be seen on the blot and the signal is not composite because the ascending arm of a Y arc cannot be seen. This data confirms that the composite signals are specific to the replication of the *redE* locus. Thus, *redE* is replicated both actively and passively in the plasmodium.
Collectively the data suggest that there are two replication origins close to redE; one upstream and one downstream. The +10' 2D gel series shows that the downstream origin is approximately 3kb downstream of the gene, while the upstream origin is approximately 8kb upstream of the gene. When both origins fire on the same chromatin, they produce replication forks that converge towards the gene and clash in the region directly 5' of redE. The existence of composite signals on fragments containing the 3' origin shows that this origin does not always fire. When the 5’ origin fires, there are two possible outcomes; (i) if the 3’ origin fires on the same chromatin, the replication forks clash in the HindIII fragment and generate a termination signal (ii) if the 3’ origin doesn’t fire, the region downstream of redE is passively replicated generating Y arcs on fragments spanning this region. Both origins fire at the onset of S-phase in the plasmodium.

5.5 DISCUSSION

The density shift analysis showed that redA is replicated after the first 45 min of S-phase. The quantitative Southern blotting analysis further refined this by showing that redA replication starts 30-60 min and finishes 120-180 min from the onset of S-phase. Most of the restriction fragments containing redA are replicated 60-90 min from the onset of S-phase. Thus, restriction fragments containing redA are replicated in mid S-phase over a broad temporal window. This pattern of replication was observed for another P. polycephalum gene; quantitative southern blotting analysis showed that the tubulin gene, altA, is replicated between 20-80 min from the onset of S-phase (Cunningham and Dove, 1993). A closer examination of the replication of altA showed that the two alleles of the gene, altA1 and altA2 are replicated at different times in S-phase; altA1 and altA2 replicate 20-40 min and 40-80 min from the onset of S-phase respectively (Cunningham and Dove, 1993). Further studies are needed to determine whether the two alleles of redA are replicated in different temporal windows. The fact that redA is replicated in mid S-phase shows that the origin(s) that result in the replication of redA are either far from the gene or late firing.

The density shift analysis showed that redB is replicated during the first 40 min of S-phase. The quantitative Southern blotting analysis further refined this by showing that redB is an early replicating gene that is replicated during the first 30 min of S-phase. The replication of redB was then examined by 2D gel electrophoresis and Southern hybridisation. This analysis pinpointed the exact replication timing of the gene; redB is replicated at the onset of S-phase in the plasmodium. The analysis also showed that two active replication origins, both of which fire at the onset of S-phase, lie either side of the gene. When these origins fire on the same chromatin strand, replication forks converge towards redB and clash in the region directly 3' of the gene. Using the measured rate of
replication fork movement through the redB locus (0.5kb/min), the two origins were placed approximately 6kb upstream and 8kb downstream of the gene.

The density shift analysis showed that redE is replicated during the first 45 min of S-phase. The replication of redE was then examined by 2D gel electrophoresis and Southern hybridisation. This analysis pinpointed the exact replication timing of the gene; redE is replicated at the onset of S-phase in the plasmodium. The analysis also showed that two replication origins, both of which fire at the onset of S-phase, are located either side of the gene. The origin that lies 3' of redE is 'weak' and does not fire 100% of the time. When both of the origins fire on the same chromatid strand, replication forks converge towards redE and clash in the region directly 5' of the gene. The origins were estimated to be 8kb upstream and 3kb downstream of the gene.

The replication timing of redA, redB and redE was examined because genes that are expressed at low levels in the plasmodium have not previously been studied. The replication timing data showed that redB and redE are replicated in a small temporal window at the onset of S-phase, while redA is replicated in a broad temporal window between 30-60 min and 90-120 min from the onset of S-phase. Thus, the replication of the red genes is not restricted to any one compartment of S-phase. The replication timing of more genes that are expressed at low levels in the plasmodium would need to be studied to see whether there is a trend in the replication timing of such genes.

Studies have shown that the upstream promoters of genes that are abundantly transcribed in the plasmodium generally contain replication origins that fire at the onset of S-phase (section 1.6.1). The replication of redB and redE was examined to see if this is true of early replicating genes that are expressed at low levels in the plasmodium. The experiments showed that redB and redE are flanked by two replication origins that fire at the onset of S-phase in the plasmodium. This arrangement is very similar to the LAV1-2 gene (section 1.6.1). Diller and Sauer (1993) showed that LAV1-2 is flanked by two 'weak' replication origins that fire at the onset of S-phase. The origins surrounding redB and redE are approximately 15kb and 11kb apart respectively, which is much smaller than the average P. polycephalum replicon size of approximately 35kb (section 1.6.1; Funderud et al., 1978a and 1979). None of the red gene promoters contained origins that fire at the onset of S-phase. Thus, the data suggest a difference in the organisation of active origins in relation to genes that are expressed at high and low levels. The replication initiation sites of more genes that are expressed at low levels in plasmodia would need to be studied to see if the organisation of active origins in relation to redB and redE is typical of genes that fall into this category.
It is possible that \textit{redB} and \textit{redE} are replicated in early S-phase because they happen to be close to abundantly transcribed genes that have origins in their promoters, i.e. the replication origins located either side of \textit{redB} and \textit{redE} could be located within the promoter regions of other genes. As not much is known about the gene density in \textit{P. polycephalum}, it is impossible to say whether it is likely that other genes would be located so close. A strategy for determining whether this is the case is outlined below. The restriction fragments containing the origins could be cloned by chromosome walking. Using the downstream \textit{redE} origin as an example, the 4.2kb \textit{ScaI} restriction fragment that stretches downstream of \textit{redE} (Fig. 3.6) could be cloned. In this procedure, a \textit{ScaI} size fractionated genomic DNA library would be made and screened by colony blotting with the \textit{redE} cDNA clone. Once the \textit{ScaI} restriction fragment was cloned, the 2.9kb \textit{BamHI KpnI} restriction fragment that is located downstream of \textit{redE} (Fig. 3.6) could be cloned. A \textit{BamHI KpnI} size fractionated genomic DNA library would be made and screened with the 3’ end of the downstream \textit{ScaI} restriction fragment. The \textit{BamHI KpnI} restriction fragment would then be analysed to see whether it contains part of a gene.
Table 5.1: Restriction mapping data for \textit{redA}, \textit{proA} and \textit{proP}

<table>
<thead>
<tr>
<th>Strain</th>
<th>M3CIV</th>
<th>M3CIV</th>
<th>M3CIV</th>
<th>TU291</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene</td>
<td>redA</td>
<td>proA</td>
<td>proP</td>
<td>proP</td>
</tr>
<tr>
<td>EcoRI</td>
<td>7.6kb</td>
<td>4kb</td>
<td>4.9kb</td>
<td>4.8kb</td>
</tr>
<tr>
<td></td>
<td>2.3kb</td>
<td>1.8kb*</td>
<td>4.8kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(I)</td>
<td></td>
<td>4.8kb</td>
<td>4.8kb</td>
</tr>
<tr>
<td>HindIII</td>
<td>-</td>
<td>-</td>
<td>4.8kb</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.6kb*</td>
<td></td>
</tr>
</tbody>
</table>

The table shows the fragment sizes obtained when a Southern blot containing \textit{EcoRI} or \textit{HindIII} digested M3CIV or TU291 genomic DNA is probed with the \textit{redA}, \textit{proA} and \textit{proP} cDNA clones. Some of the probes hybridised to two restriction fragments; either there is an allelic polymorphism at the gene locus (marked as *) or the gene contains an internal enzyme restriction site (marked as I).
Figure 5.1: Restriction map of the redB locus

All of the restriction sites were mapped in the diploid strain, M3CIV. Allelic polymorphisms are marked *. The sequential order of restriction sites which are close together may not be correct because the map was constructed with the data obtained from several Southern blots. The region marked redB represents that covered by the cDNA and partial genomic clone (Bailey et al., 1999). The internal SacI and PstI sites are not in the known redB sequence and have been positioned by mapping. Presumably these sites are located in unsequenced introns.
Figure 5.2: Density shift experiments

Schematic diagrams showing the procedure for density shift analysis. (a) At the start of S-phase, a plate grown macroplasmodium is placed onto filter paper soaked with BrdU. BrdU molecules (*) are taken into the plasmodium and incorporated into the replicating DNA until the plasmodium is harvested for DNA isolation. (b) The isolated DNA molecules have replicated regions, that are BrdU substituted, interspersed with non-replicated regions, that are non-BrdU substituted. In order to reduce the number of part-substituted DNA molecules, the DNA is digested with a restriction enzyme. (c) Most of the smaller DNA molecules are now either BrdU-substituted or non-BrdU substituted. The digested DNA is then centrifuged on a caesium chloride (CsCl) gradient to separate the normal, LL DNA from the density shifted, HL DNA. (d) After centrifugation, the LL and the HL DNA fractions can be easily distinguished. (e) They are isolated separately from the centrifuge tube with a syringe. (f) The two DNA fractions are now separate.
Figure 5.3: The separation of LL and HL DNA on a CsCl gradient

The DNA is stained with ethidium bromide and the outline of the centrifuge tube has been drawn onto the photograph in white. DNA that was trapped in the neck of the tube and has not spun down properly can be seen at the top of the tube. The LL DNA fraction is more abundant than the HL DNA fraction because the DNA came from a plasmodium that was treated with BrdU for 45 mins; at this point in S-phase about one third of the genome has been replicated.
Figure 5.4: Density shift analysis of the replication timing of *redA* and *redB*

Genomic DNA was isolated from an M3CIV macroplasmodium in G2-phase of the cell cycle and from a macroplasmodium that had been treated with BrdU for the first 40 min of S-phase (section 2.5.2). The isolated DNAs were digested with *Eco*RI (section 2.6.1.3). The DNA from the BrdU-treated plasmodium was then run on a CsCl gradient to separate the LL and the HL DNA (section 2.7.2). Approximately 5μg of the G2-phase DNA and the LL and the HL DNA was run on a 0.8% agarose gel with λ *Hind*III and φX174 *Hae*III DNA size markers (section 2.6.2). The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The cDNA clones for *redA* (Bailey *et al.*, 1999) and the *proA* (Binette *et al.*, 1990) were radio-labelled and hybridised to the blot (section 2.7.4). The blot was exposed to x-ray film at −80°C with two intensifying screens for 1-5 days.

(a) a photograph of the ethidium bromide stained gel under UV illumination.
(b) autoradiograph of the blot after exposure at −80°C.

The procedure was repeated using *Hind*III digested LL and the HL DNA from a plasmodium that had been treated with BrdU for the first 40 min of S-phase. The resulting blot was probed with the *redB* (Bailey *et al.*, 1999) and the *proP* (Binette *et al.*, 1990) cDNA clones.

(c) a photograph of the ethidium bromide stained gel under UV illumination.
(d) autoradiograph of the blot after exposure at −80°C.
Figure 5.5: Density shift analysis of the replication timing of redA, redB and redE

Genomic DNA was isolated from a TU291 macroplasmodium in G2-phase of the cell cycle and from a macroplasmodium that had been treated with BrdU for the first 45 min of S-phase (section 2.5.2). The isolated DNAs were digested with EcoRI (section 2.6.1.3). The DNA from the BrdU-treated plasmodium was then run on a CsCl gradient to separate the LL and the HL DNA (section 2.7.2). Approximately 5μg of the G2-phase and the LL and the HL DNA was run on a 0.8% agarose gel with λ HindIII and φX174 HaeIII DNA size markers (section 2.6.2). The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The cDNA clones for redE and the proP (Binette et al., 1990) were radio-labelled and hybridised to the blot (section 2.7.4). The blot was exposed to x-ray film at -80°C with two intensifying screens for 1-5 days.

(a) a photograph of the ethidium bromide stained gel under UV illumination.
(b) autoradiograph of the blot after exposure at -80°C.

The procedure was repeated using EcoRI digested DNA from an M3CIV plasmodium that had been treated with BrdU for the first 40 min of S-phase. The resulting blot was hybridised to a probe mixture containing the redA, redB (Bailey et al., 1999), redE and the proP (Binette et al., 1990) cDNA clones.

(c) a photograph of the ethidium bromide stained gel under UV illumination.
(d) autoradiograph of the blot after exposure at -80°C.
Figure 5.6: Quantitative Southern blot analysis of the replication timing of redA and redB

Genomic DNA was isolated from an M3CIV plasmodium in G2-phase and from plasmodia 30 min, 60 min, 90 min, 120 min, 180 min and 300 min from the onset of S-phase (sections 2.5.1 and 2.5.3). 5μg of each DNA sample was digested with EcoRI (section 2.6.1.2) and electrophoresed with λ HindIII DNA size markers on a 0.8% agarose gel (section 2.6.2). The DNA was stained with ethidium bromide, transferred to membrane and fixed to the membrane by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1).

The cDNA clones for redA, redB (Bailey et al., 1999) and proP (Binette et al., 1990) were radio-labelled and hybridised to the blot (section 2.7.4). The blot was exposed to x-ray film at -80°C with two intensifying screens for approximately 3 days.

(a) photograph of the ethidium bromide stained gel under UV illumination.

(b) autoradiograph of the blot after exposure at -80°C.
The intensity of each band on the Southern blot in Figure 5.6 was quantitatively measured to give a 'volume report' in arbitrary units. The G2 volume reports were called 1.0 because restriction fragments containing redA, redB and proP are equally abundant in this sample. The volume reports in all other lanes were then normalised to the G2 volume reports. The normalised volume reports were compared to give band ratios. In this procedure, the normalised volume report of one band was divided by the normalised volume report of another band for every combination of genes within each gel lane. This gave three sets of band ratios:

(a) shows the proP/redA band ratios.
(b) shows the proP/redB band ratios.
(c) shows the redB/redA band ratios.
Schematic diagrams showing the procedure for 2D gel electrophoresis. (a) DNA is extracted from a plasmodium in S-phase, imbedded in an agarose plug and digested with a restriction enzyme. (b) The DNA in the agarose plug is then run on a 0.4% agarose gel. (c) After running, the gel lane containing the size fractionated genomic DNA is excised from the gel and placed horizontally across the top of a gel bed, in the position the lanes usually occupy. (d) Melted 1.1% agarose is poured around the 0.4% gel slice to form a new gel, which is run under conditions of high voltage. After running, the DNA is transferred to membrane and single-copy restriction fragments are located by Southern hybridisation. The linear DNA molecules form a continuous diagonal curve on a 2D gel; (f) the small DNA molecules run far into the gel, and (e) the larger DNA molecules run less far into the gel. The replication intermediates run more slowly than linear DNA in the first and second dimensions and are located above the line of linear DNA.
Figure 5.9: Interpreting 2D gels

Schematic diagrams showing the 2D gel patterns generated by four types of replication intermediates (based on Brewer and Fangman, 1987; Diller and Sauer, 1993). Each diagram depicts the replication of a hypothetical 3kb restriction fragment. In the lower panels, the position that linear DNA runs to is indicated by a straight dashed line and the expected migration of the replication intermediates is shown above. Linear copies of the hypothetical restriction fragment run at the line of linear DNA. The point at which the restriction fragment touches this line is called the 1x spot (marked o, 3kb in each panel). When the replication of the restriction fragment is almost complete the fragment is approximately 6kb and sometimes runs at the line of linear DNA. The point at which the restriction fragment touches this line is called the 2x spot (marked o, 6kb in each panel).

(a) a simple Y arc.
(b) a bubble arc.
(c) a bubble to Y transition.
(d) a double Y.

The position of a simple Y arc is indicated by a dashed curve in panels (b), (c) and (d). The pattern shown in (d) is observed when two replication forks meet exactly in the centre of a restriction fragment. If one replication fork enters the restriction fragment before the other the replication forks would collide off-centre somewhere along the X-spike (x); forks that collide slightly off-centre are found towards the top of the spike, while forks that collide very off-centre are found towards the bottom of the spike.
Figure 5.10: 2D gel analysis of the replication of \textit{redB}

2D gels containing approximately 10μg of digested genomic DNA were prepared as described in section 2.7.3. The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The \textit{redB} cDNA (Bailey et al., 1999) was radio-labelled and hybridised to the blots (section 2.7.4). The blots were exposed to x-ray film at -80°C with two intensifying screens for 1-5 days.

The replication of the \textit{PstI} fragments was examined at +4', +10' and +15' from the onset of S-phase. The upstream and the downstream \textit{PstI} fragments are named A and B respectively. The two schematic drawings depict the movement of replication forks through A and B at +4', +10' and +15' for (a) two replication forks diverging from \textit{redB}, and (b) two replication forks converging towards \textit{redB}. In these drawings – indicates unreplicated DNA, = indicates replicated DNA, <<<< or >>> indicates the region spanned by replication forks and <=> indicates a replication bubble.

The replication of the \textit{HindIII KpnI} fragment was examined at +4'. If replication forks diverge from \textit{redB} the \textit{HindIII KpnI} fragment will be replicating at +4' [\textbf{prediction (a)}]. If replication forks converge towards \textit{redB}, the \textit{HindIII KpnI} fragment will not be replicating at +4' [\textbf{prediction (b)}]. There is a 1x spot on the blot (the second minor spot on the diagonal of linear molecules is a partial digest). The absence of arcs shows that there are no replication forks within the gene region at this time.
Predicted either (a) or (b) observed
Figure 5.11: Further 2D gel analysis of the replication of redB

2D gels containing approximately 10µg of digested genomic DNA were prepared as described in section 2.7.3. The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The redB cDNA (Bailey et al., 1999) was radio-labelled and hybridised to the blots (section 2.7.4). The blots were exposed at -80°C with two intensifying screens for 1-5 days.

The replication of the HincII and Xhol restriction fragments was examined at +4', +10' and +15' from the onset of S-phase. Termination signals (T) can be seen on the +15' HincII and Xhol blots. The top drawing depicts the movement of replication forks through the upstream and the downstream PstI fragments at +4', +10' and +15' as previously determined (see Fig. 5.10). The other two drawings show the movement of replication forks through the HincII and the Xhol restriction fragments at +4', +10' and +15'. In the schematic drawings – indicates unreplicated DNA, = indicates replicated DNA and <<<< or >>> indicates the region spanned by replication forks.
Figure 5.12: 2D gel analysis of the replication of redE

2D gels containing approximately 10μg of digested genomic DNA were prepared as described in section 2.7.3. The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The redE cDNA was radio-labelled and hybridised to the blots (section 2.7.4). The blots were exposed to x-ray film at -80°C with two intensifying screens for 1-5 days.

The replication of the BglII, KpnI and EcoRV restriction fragments was examined +10' into S-phase and the replication of the HindIII and EcoRI fragments was examined +15' into S-phase. Termination signals (T) can be seen on the +15' HindIII and EcoRI blots. The diagram depicts the movement of replication forks through the region spanned by the EcoRI and HindIII restriction fragments at +15'. In this drawing – indicates unreplicated DNA, = indicates replicated DNA and <<< or >>> indicates the region spanned by replication forks.
2D gels containing approximately 10μg of digested genomic DNA were prepared as described in section 2.7.3. The DNA was transferred to membrane and fixed by baking at 80°C in a vacuum oven for 1 hour (section 2.7.1). The redE and proP (Binette et al., 1990) cDNA clones were radio-labelled and hybridised to the relevant blots (section 2.7.4). The blots were exposed to x-ray film at -80°C with two intensifying screens for 1-5 days.

(a) the replication of the KpnI fragment containing redE and its 3' origin at +10'.
(b) the replication of a restriction fragment containing proP and its 5' origin at +10'.
(c) the replication of the KpnI restriction fragment containing redE and its 3' origin at +5'.
(d) the replication of a restriction fragment containing proP and its 5' origin at +5'.
CHAPTER 6: THE DEVELOPMENT OF A LUCIFERASE REPORTER GENE SYSTEM

6.1 INTRODUCTION

Hemon (1996) wanted to make a reporter gene construct for stable transformations that could be modified to allow her to examine the activity of various promoters and the function of various genes in *P. polycephalum*. Such a construct needed to contain a reporter gene in addition to a selectable marker, to allow the selection of stable transformants. Since the reporter gene *luc*, and the selectable marker *hph*, had been used in *P. polycephalum* (sections 1.4.1 and 1.4.2; Burland *et al.*, 1993; Bailey *et al.*, 1994), Hemon (1996) decided that her construct should contain *PardB-luc-hph-PardC*. The activity of a promoter could be examined by placing *luc* under the control of a promoter from a gene of interest, e.g. *PardB* could be replaced with a *red* gene promoter. Any stable transformants would be selected on the basis of *HygR* and the activity of the promoter could be examined throughout development by assaying the transformant cells for Luc activity. The function of a gene of interest could be examined by mis-expressing the gene, e.g *luc* could be replaced with a *red* gene. Any stable transformants would be selected on the basis of *HygR* and phenotype of the transformant cells could be examined throughout development by time-lapse filming or immunofluorescence microscopy.

Hemon (1996) made such a construct by cloning the 3.6kb *PardB-luc* cassette from pFH1 (Fig. 6.1a) into pTB37 which contained the 2.4kb *PardC-hph* cassette. The resulting plasmid, pFH3, contained *PardB-luc-hph-PardC* such that the *luc* and *hph* genes were transcribed towards each other (Fig. 6.1b). The transcriptional terminator sequence, *TardC*, was not included in the vector because Burland *et al.* (1991) had shown that its presence did not increase the expression of the reporter gene, *cat* (Hemon, 1996).

Hemon (1996) carried out several transient transformation experiments to compare the *luc* expression of pFH1 and pFH3 to that of p*PardB-luc* which is known to express *luc* well (Bailey *et al.*, 1994). Axenically grown LU352 amoebae were aliquoted at a density of approximately $5 \times 10^7$ cells per tube. One of the three *luc* plasmids was added to an aliquot of cells and the cell/DNA mixtures were electroporated at the optimum electroporation settings (section 1.4.1). The cells were transformed with 1μg of a 6.6kb plasmid and the weight was adjusted for bigger plasmids so that each aliquot of cells was transformed with approximately the same number of plasmid molecules. Approximately 2.5 hours after electroporation, $1 \times 10^7$ cells were removed from each sample were assayed for Luc activity. Cells that were transformed with p*PardB-luc* had the highest levels of *luc* expression followed by pFH1 and pFH3 (Hemon, 1996). On average, cells that were transformed with
pFH1 had 28% of the Luc activity of cells that were transformed with \textit{pPardB-luc}, while cells that were transformed with pFH3 had just 19% of the Luc activity of cells that were transformed with \textit{pPardB-luc}. Since pFH3 cells were expressing \textit{luc}, Hernon (1996) went ahead with stable transformation experiments.

Burland and Pallotta (1995) showed that (i) linear DNA integrates into the genome more frequently than circular DNA and (ii) smaller pieces of DNA integrate into the genome more frequently than larger pieces of DNA (section 1.4.2). Thus, Hernon (1996) digested pFH3 with \textit{NsiI} and \textit{SfiI} which gave the \textit{PardB-luc-hph-PardC} cassette on a 6kb restriction fragment (Fig. 6.1b). LU352 amoebae were transformed by electroporation with digested pFH3 DNA and plated onto plates containing Hyg. One of the transformation experiments produced two Hyg\textsuperscript{R} transformants, FH3T1 and FH3T3 (Hernon, 1996). The transformant amoebae were inoculated into axenic culture and, when the amoebae were growing well, approximately 1x10\textsuperscript{7} cells were assayed for Luc activity. Surprisingly, no Luc activity was detected in either of the transformant strains. To look more closely at what was happening, genomic DNA and total RNA was isolated from the transformant cell lines.

Southern blotting analysis showed that both \textit{hph} and \textit{luc} were integrated into the nuclear genomes of FH3T1 and FH3T2. Digested genomic DNA from FH3T1 and FH3T2 produced identical bands on the Southern blots, suggesting that the transformants are siblings that arose by the division of a single stably transformed cell (data not shown; Hernon, 1996). Northern blotting analysis, using amoebal and plasmodial RNA from both transformant strains, produced some unexpected results. The \textit{luc} transcript is usually 1700nt in length, but in the transformant cell lines the \textit{luc} probe hybridised to an mRNA that was 5100nt in length (data not shown; Hernon, 1996). The mRNA had an additional 3400nt of 3' untranslated region which showed that the transcriptional termination of the \textit{luc} transcript was not occurring properly. As the \textit{luc}, \textit{hph} and \textit{PardC} elements have a total length of 4.1kb, the transcriptional termination of the \textit{luc} mRNA was occurring approximately 1kb outside of the integrated construct (Hernon, 1996). The \textit{hph} transcript is usually 1300nt in length, but in the transformant cell lines the \textit{hph} probe hybridised to an mRNA that was 1900nt in length (data not shown; Hernon, 1996). The mRNA carried an additional 600nt of 3' untranslated region which showed that the transcriptional termination of the \textit{hph} transcript was not occurring properly. As the \textit{hph} and \textit{luc} elements have a total length of 3kb, the transcriptional termination of the \textit{hph} transcript was occurring approximately 600bp into the \textit{luc} gene (Hernon, 1996). These findings explained the lack of Luc activity in the transformant cells. Hernon (1996) proposed that the elongated \textit{luc} mRNA was not being processed properly and that this affected the production of the Luc protein (Hernon, 1996). The elongated \textit{hph} transcript did not
significantly affect the production of the Hyg protein because FH3T1 and FH3T3 transformant cells were HygR.

The first aim of this work was to modify pFH3 to ensure that the transcriptional termination of the luc and hph transcripts occurs correctly. The second aim was to indirectly investigate the function of red gene(s) by using the construct to mis-express the gene(s) and analysing the phenotype of any stable transformants.

RESULTS

6.2 VECTOR CONSTRUCTION

TardC is a 600bp A-T rich terminator sequence from the 3' end of the actin gene, ardB (Burland et al., 1991). It was thought that inserting TardC in between the luc and hph genes to form a construct containing PardB-luc-TardC-hph-PardC would ensure the correct transcriptional termination of the luc and hph genes. Such a construct was made using two plasmids: (i) pPardB-luc which contains the 3.6kb cassette, PardB-luc, cloned into pGem-7Z (Promega; Fig. 6.2a) and (ii) pTB40 which contains the 3kb cassette, PardC-hph-TardC, cloned into the same vector (Fig. 6.2b).

Approximately 30μg of pTB40 was digested with HindIII and KpnI to release the 3kb cassette, PardC-hph-TardC (Fig. 6.2b; section 2.3.1.1). The 3kb restriction fragment was then gel purified and its ends were blunted with the Klenow fragment of E. coli DNA polymerase before resuspension in a volume of 20μl (sections 2.3.3 and 2.3.4). Meanwhile, 15μg of pPardB-luc was linearised with StuI (Fig. 6.2a) and treated with alkaline phosphatase before resuspension in a volume of 20μl (sections 2.3.3 and 2.3.4). Two ligations were set up: (i) 0.5μl of vector (linear pTB40) was ligated to 1μl of insert (PardC-hph-TardC) (ii) 0.5μl of vector was ligated to 2.5μl of the insert (section 2.3.5). The ligations had a final volume of 10μl. Approximately 3μl of each ligation was transformed into E. coli XL1-Blue and a proportion of each ligation was plated onto LB-amp plates (section 2.3.5). Hundreds of colonies could be seen on the plates after overnight incubation. Twenty colonies were chosen at random, and plasmid DNA was extracted from each (section 2.2.6). Each plasmid DNA sample was then divided into two; half was digested with EcoRI and half was digested with BamHI (section 2.3.1.2). The digests were later run on a 0.8% agarose gel (data not shown; section 2.3.2). Two different plasmids were expected from the ligations because the blunt-ended PardC-hph-TardC cassette could integrate into the blunt-ended pPardB-luc vector in either orientation (Fig. 6.2). The EcoRI and BamHI digest patterns that each plasmid would give were predicted. The electrophoresis showed that approximately half of the isolated
plasmids gave the predicted digest patterns of a plasmid containing \textit{PardB-luc-TardC-hph-PardC} such that the \textit{luc} and \textit{hph} genes were transcribed in opposite directions, towards \textit{TardC}. While, the other half gave the predicted digest patterns of a plasmid containing \textit{PardB-luc-PardC-hph-TardC} such that the \textit{luc} and \textit{hph} genes were transcribed in the same direction, towards \textit{TardC}. These plasmids were named pES1 and pES2 respectively (Fig. 6.2c and d) and frozen stocks were made of colonies that contained each of the plasmids (section 2.3.6).

6.3 TRANSIENT TRANSFORMATIONS

Five transient transformation experiments were carried out to compare the \textit{luc} expression of pES1, pES2 to that of pFH1, pFH3, \textit{pPardB-luc} and \textit{pPardC-luc}. Axenically grown LU352 amoebae were aliquoted into seven tubes at a density of approximately $5 \times 10^7$ cells per tube (section 2.1.4). One of the six \textit{luc} plasmids was added to an aliquot of cells and these cell/DNA mixtures and the remaining aliquot of cells were electroporated at the optimum electroporation settings (section 2.1.5.1). Approximately 2.5 hours after electroporation, $1 \times 10^7$ cells from each sample were assayed for luciferase activity using a luminometer, see below. The cells were transformed with 1$\mu$g of a 6.6kb plasmid and the weight was adjusted for bigger plasmids so that each aliquot of cells was transformed with approximately the same number of plasmid molecules, e.g 1.43$\mu$g of a 9.6kb plasmid was transformed. The cells that were electroporated with no DNA determined the background light output of the cells.

There were two steps in the analysis of the data from a single transformation experiment:

(i) The light output of $1 \times 10^7$ cells from each sample was measured using a luminometer. The luminometer measured light emission in relative light units (RLU) over a period of 30 seconds (data not shown; section 2.1.5.1). There was a positive correlation between the light output and RLU reading; the higher the light output, the higher the RLU reading.

(ii) The \textit{pPardB-luc} RLU reading was called 100% and the remaining five RLU readings were normalised to this reading, e.g in transformation 1 the \textit{pPardB-luc} had an RLU reading of 519730 and pFH1 had an RLU reading of 643870. Thus, the normalised pFH1 RLU reading is 124% ($643870 \div 519730 \times 100$). A second data set was created by calling the \textit{pPardC-luc} RLU reading 100% and normalising the remaining five RLU readings to this reading, as before. Both data sets can be seen in Table 6.1.

The RLU readings of each experiment were normalised to the RLU readings of both \textit{pPardB-luc} or \textit{pPardC-luc} so that the \textit{luc} expression of each plasmid could be easily compared.
The normalised pPardB-luc RLU readings from the five experiments were averaged and plotted on a graph (Fig. 6.3a). On average, cells that were transformed with pES1 had the highest levels of luc expression followed by pES2, pPardB-luc, pFH1, pPardC-luc and pFH3. The same hierarchy of luc expression was seen when the normalised pPardC-luc RLU readings from the five experiments were averaged and plotted on a graph (Fig. 6.3b). Cells that were transformed with pES1 or pES2 expressed luc well in comparison to cells that were transformed with any other plasmid. This was not surprising for pES1 which should produce normal luc transcripts because TardC is located directly downstream of luc (Fig. 6.2c). However, this was surprising for pES2 which may produce elongated luc transcripts because TardC is located some distance from luc (Fig. 6.2d). It is possible that pES2 cells have slightly lower levels of luc expression than pES1 cells because the luc transcript is elongated in these cells.

Cells that were transformed with pPardB-luc had higher levels of luc expression than cells that were transformed with pPardC-luc. This finding contradicts the work of Bailey et al. (1994) which showed that cells which were transformed with pPardC-luc had higher levels of luc expression than cells that were transformed with pPardB-luc. Although the authors speculate that it was not a fair test because cells were transformed with 1µg of each plasmid, and as pPardC-luc is 500bp larger than pPardB-luc, the cells were therefore transformed with a greater number of pPardC-luc plasmid molecules (Bailey et al., 1994). The data for pPardB-luc, pFH1 and pFH3 is consistent with the work of Hemon (1996) which showed that cells that were transformed with pPardB-luc had the highest levels of luc expression followed by pFH1 and pFH3. Cells that were transformed with pFH3 consistently had the lowest levels of luc expression. This is not surprising since Hemon (1996) showed that pFH3 stable transformants had very elongated luc transcripts.

6.4 STABLE TRANSFORMATIONS

Both pES1 and pES2 were used for stable transformation experiments because the transient transformation experiments showed that the vectors expressed luc well. The only difference between pES1 and pES2 is the orientation of PardC-hph-TardC. Although pES2 is not the vector of choice, it was included in the transformations because if pES1 stable transformants produced a normal luc mRNA, and pES2 stable transformants produced an elongated luc mRNA, this would show that TardC is essential for the correct transcriptional termination of luc. The analysis of the pFH3 stable transformants, FH3T1 and FH3T3 suggested that the transcriptional termination of the luc transcript occurred in the surrounding DNA, approximately 1kb outside of the integrated construct (section 6.1). It was hypothesised that in cells that were stably transformed with pFH3, the exact length of the luc transcript would depend on the site of integration. To see if this was true, and
pFH3 transformants have differently sized luc transcripts, pFH3 was included in the stable transformation experiments.

To make the transforming DNA as small as possible (section 6.1), the sequence of the pES1 and pES2 original host vector, pGem-7Z (Promega) was examined to see what restriction enzymes would release the PardB-luc-TardC-hph-PardC cassette of pES1 and the PardB-luc-PardC-hph-TardC cassette of pES2. The smallest restriction fragment that could be easily obtained for both vectors was the 8.8kb ApaI SspI fragments which contained the 6.6kb PardB-luc-TardC-hph-PardB or PardB-luc-PardC-hph-TardC cassettes (Fig. 6.2c and d). As before, the 6kb PardB-luc-hph-PardC cassette was excised from pFH3 with SfiI and NsiI (Hernon, 1996; Fig. 6.1b).

Approximately 50µg of pES1, pES2 and pFH3 was digested with the appropriate restriction enzymes before resuspension in a volume of 40µl (section 2.3.2). A small amount of each digested plasmid was run on a 0.8% agarose gel with λ HindIII and φx HaeIII DNA size markers, stained with ethidium bromide and visualised under UV illumination (section 2.3.2). The concentration of the digested plasmid DNA was estimated by comparing the plasmid DNA and the marker DNA band intensities. The vector bands were not removed from the required construct DNA because these restriction fragments do not contain the hph gene and cells that become stably transformed with these fragments will not be selected.

Aliquots of axenically grown LU352 amoebae were mixed with approximately 3µg each digested luc plasmid and electroporated at the optimum electroporation settings (section 2.1.5.2). The cells were then plated onto plates containing hyg and incubated at 30°C for up to four weeks (section 2.1.5.2). Every few days, the plates were checked for signs of colony growth. From a total of eight stable transformation experiments, one pFH3 HygR stable transformant (FH3T3), one pES1 HygR stable transformant (ES1T1) and no pES2 transformants were identified. Since eight transformation experiments were conducted, pFH3 and pES1 yielded transformants at a frequency of 4x10⁻⁸ cell⁻¹.

6.5 THE ANALYSIS OF FH3T3 AND ES1T1

The FH3T3 and the ES1T1 amoebal colonies were sub-cultured onto fresh selective media and grown to confluence before they were sub-cultured onto three fresh plates (section 2.1.5.2). The cells on the first plate were used to make a frozen stock of the transformant strain (section 2.1.3), the cells on the second were transferred to axenic culture (section 2.1.4), and the cells on the third plate were induced to develop into apogamic plasmodia (section 2.1.6). A plasmodium of each strain was later transferred to axenic culture (section
Total RNA and genomic DNA was isolated from axenically grown amoebae and microplasmodia (sections 2.2.2, 2.2.3 and 2.2.4). FH3T1 amoebae (Hemon, 1996) were grown from cysts and subjected to the same procedures as the FH3T3 and ES1T1 cells, so that all three transformants could be compared in the same experiments.

Approximately $1 \times 10^7$ cells were removed from axenically grown FH3T1, FH3T3, ES1T1 and LU352 amoebal cultures and assayed for Luc activity (section 2.1.5.1). The RLU reading of the LU352 amoebae gave the background light output of the cells. Thus, this reading was deducted from the remaining three readings before they were analysed. The FH3T1 and the ES1T1 cells did not have any Luc activity. The FH3T3 cells had significant levels of Luc activity (data not shown). The data are further discussed in section 6.6.

Northern blotting was used to look at the expression of hph and luc in the transformant cell lines. Two northern blots were made which contained 10μg of total RNA from (i) FH3T1 amoebae (ii) FH3T1 microplasmodia (iii) FH3T3 amoebae (iv) FH3T3 microplasmodia (v) ES1T1 amoebae and (vi) ES1T1 microplasmodia. The first northern blot was probed with PardB-luc (data not shown) and the second northern blot was probed with PardC-hph (data not shown). The blots were exposed to x-ray film at $-80^\circ$C for 2 weeks with an intensifying screen. When the x-ray films were developed, no bands were visible (data not shown) which indicated a problem with the RNA samples. These experiments were not repeated due to time constraints.

Southern blotting was used to determine whether integration events had taken place in the transformant cell lines. A Southern blot containing 3μg of HindIII digested DNA from LU352, FH3T1, FH3T3 and ES1T1 microplasmodia was probed with PardC-hph (Fig. 6.4). The PardC-hph probe will hybridise to the endogenous copy of PardC as well as to any integrated construct(s). The LU352 lane contains a single band of approximately 7.5kb which indicates the endogenous copy of PardC is contained on a 7.5kb HindIII restriction fragment. The FH3T1 DNA contains the wild-type PardC band in addition to a band of approximately 9.4kb which indicates an integration event has taken place. This is consistent with previous findings (Hemon, 1996). The FH3T3 lane contains the wild-type PardC band in addition to two bands of approximately 3kb and 2kb which indicates that integration event(s) have taken place. The ES1T1 lane contains the wild-type PardC band in addition to a band of approximately 20kb which could mean one of two things: either (i) an integration event has taken place or (ii) some of the genomic DNA is undigested and contains wild-type PardC. More extensive Southern blotting is required to characterise the integration events that have taken place.
6.6 DISCUSSION

Hemon (1996) made a reporter gene construct, pFH3, for stable transformations. The plasmid contained PardB-luc-hph-PardC, such that the luc and hph genes were transcribed in opposite directions, towards each other. She transformed amoebae with digested pFH3 and obtained two stable transformants, FH3T1 and FH3T2 (Hemon, 1996). Surprisingly, FH1T1 and FH3T2 amoebae had no Luc activity. Southern blotting analysis showed that the transformants FH1T1 and FH3T2 were siblings that had arisen by the division of a single transformed cell and that both had a single integrated copy of the pFH3 fragment. Northern blotting showed that the transformants contained an elongated luc mRNA of approximately 5100nt. The transcription of the luc transcript passed through luc (1.1kb) and continued through hph (1.1kb) and PardC (1.3kb) until transcriptional termination occurred approximately 1kb outside of the integrated construct (Hemon, 1996). The analysis also showed that the transformants contained an elongated hph mRNA of approximately 1900nt. The transcription of the hph transcript passed through hph (1.3kb) and continued through luc until transcriptional termination occurred approximately 600bp into the luc gene (Hemon, 1996). These findings explained the lack of Luc activity in the transformant cells. Hemon (1996) proposed that the elongated luc mRNA was not being processed properly and that this affected the production of the Luc protein. The elongated hph transcript did not significantly affect the production of the Hyg protein because FH3T1 and FH3T2 transformant cells were HygR (Hemon, 1996).

TardC is a 600bp A-T rich terminator sequence from the 3' end of the actin gene, ardC (Burland et al., 1991). It was thought that inserting TardC in between the luc and hph genes to form a construct containing PardB-luc-TardC-hph-PardC would ensure the correct transcriptional termination of the luc and hph genes. I made two constructs from pPardB-luc and pTB40. The first construct, pES1, contained PardB-luc-TardC-hph-PardC such that the luc and hph genes were transcribed in opposite directions, towards TardC. While, the second construct, pES2, contained PardB-luc-TardC-hph-PardC such that the luc and hph genes were transcribed in the same direction, towards TardC. Transient transformation experiments were used to compare the luc expression of pES1 and pES2 to that of pPardB-luc, pPardC-luc, pFH1 and pFH3. These experiments showed that cells that were transformed with pES1 had the highest levels of luc expression followed by pES2, pPardB-luc, pFH1, pPardC-luc and pFH3.

Stable transformation experiments were carried out using pES1 and pES2 because the transient transformation experiments showed that both of the plasmids expressed luc well. Although pES2 was not the vector of choice, it was included in the transformations as a
control to see whether the correct transcriptional termination of the \textit{luc} transcript is dependent on the presence of \textit{TardC}. Hemon (1996) showed that the transcriptional termination of the \textit{luc} transcript occurred outside of the integrated construct in FH3T1 and FH3T3. pFH3 was included in the stable transformations to see whether the exact length of the \textit{luc} transcript was dependent on the site of integration for this vector. The plasmids pES1 and pES2 were digested with \textit{ApaI SspI} to release the 8.8kb fragment containing the \textit{PardB-luc-TardC-hph-PardC} or the \textit{PardB-luc-PardC-hph-TardC} cassettes. While, pFH3 was digested with \textit{NsiI SalI} to release the 6.6kb \textit{PardB-luc-hph-PardC} cassette. The cells were transformed with approximately 3\textmu g of plasmid DNA and stable transformants were selected on the basis of \textit{Hyg}^R. A total of eight transformation experiments yielded one pFH3 \textit{Hyg}^R stable transformant, FH3T3, one pES1 \textit{Hyg}^R stable transformant, ES1T1, and no pES2 stable transformants. It was calculated that the pFH3 and pES1 fragments yielded transformants at a frequency of $4 \times 10^{-8}$ per cell. This is within the range that has been previously observed, since various transformation constructs have yielded stable transformants at a frequency of between $1.8 \times 10^{-7}$-$1 \times 10^{-8}$ per cell (Burland \textit{et al.}, 1993; Burland and Pallotta, 1995). The transformants were then analysed by Luc assays and northern and Southern blotting. FH3T1 (Hemon, 1996) was included in the analysis as a comparison.

As expected, FH3T1 amoebae had no Luc activity (Hemon, 1996). FH3T3 amoebae had significant levels of Luc activity, while ES1T1 amoebae had no Luc activity. The work of Hemon (1996) suggests that pFH3 transformants will have an elongated \textit{luc} transcript of at least 4100nt because the transcriptional termination of this mRNA does not occur within the integrated DNA. It is surprising that FH3T3 amoebae have significant levels of Luc activity because these cells are expected to contain an elongated \textit{luc} transcript that could affect the production of functional Luc protein. It is also surprising that ES1T1 amoebae have no Luc activity because it was thought that the presence of \textit{TardC} would ensure the correct termination of the \textit{luc} transcript, leading to the production of functional Luc protein. Northern blotting was carried out to determine the \textit{luc} and \textit{hph} transcript sizes in the transformant strains, but failed because the RNA samples were degraded. The work was not repeated due to time constraints.

Southern blotting showed that there had been an integration event in FH3T1 and FH3T3. It was unclear whether an integration event had occurred in ES1T1. Pierron \textit{et al.} (1999) showed that the transforming DNA is sometimes only partially integrated into the nuclear genome of stable transformants. The authors examined a \textit{Hyg}^R transformant that had been transformed with an 8.8kb restriction fragment containing \textit{PardC-hph-TardC}. Southern blotting analysis showed that only 2.1kb of the 8.8kb construct had been integrated into the
nuclear genome at a single integration site. Subsequent sequence analysis showed that while *PardC* was fully integrated, *hph* was truncated and *TardC* and other vector sequences were missing. The truncated *hph* gene encoded a Hyg protein that was missing 18 amino acids. But despite being truncated, the Hyg protein was still functional since the transformant was selected on the basis of Hyg resistance. One possible explanation for the lack of Luc activity in ES1T1 transformant cells is that the 8.8kb *PardB-luc-TardC-hph-PardC* cassette was only partially integrated, so that the cells are Hyg<sup>k</sup> but lack a functional copy of *luc*. The integration event needs to be further characterised to see if this is the case.

The work described in this chapter is unfinished. To complete the work, total RNA needs to be isolated from the transformant strains and analysed by northern blotting to determine the sizes of the *luc* and *hph* transcripts.
Table 6.1: The relative Luc activity of pPardB-luc, pPardC-luc, pFH1, pFH3, pES1 and pES2 in transient transformation experiments

(a) For each of the five experiments the pPardB-luc relative light output reading was called 100%, and the relative light output readings for the other five plasmids were normalised to the relative light output of pPardB-luc.

(b) For the same five experiments, the pPardC-luc relative light output reading was called 100%. The relative light output readings for the other five plasmids were then normalised to the relative light output of pPardC-luc.

The Luc activity of six luc bearing plasmids (pPardB-luc, pPardC-luc, pFH1, pFH3, pES1 and pES2) was determined in five independent transient transformation experiments (# 1-5). In this procedure, the plasmids were transformed into amoebae by electroporation (section 2.) and the cells were later assayed for luciferase activity (section 2.). The luciferase assay gave each aliquot of cells a measure of light output in ‘relative light units’ (data not shown).
Figure 6.1: The structure of pFH1 and pFH3

(a) pFH1
- **PardB**: 6800 bp
- **luc**: 6800 bp

(b) pFH3
- **PardB**: 9300 bp
- **luc**: 9300 bp
- **PardC**: 9300 bp
- **hph**: 9300 bp
- **NsiI**: 9300 bp
- **SfiI**: 9300 bp
Figure 6.2: The structure of \textit{pPardB-luc}, \textit{pTB40}, \textit{pES1} and \textit{pES2}

(a) p\textit{PardB-luc} 6000 bp

(b) p\textit{TB40} 6000 bp

(c) p\textit{ES1} 9600 bp

(d) p\textit{ES2} 9600 bp
Figure 6.3: The average relative levels of Luc activity of pPardC-luc, pPardC-luc, pFH1, pFH3, pES1 and pES2 in transient transformation experiments

(a) shows the relative levels of Luc activity, normalised to pPardB-luc, for the five transient transformation experiments (see Table 6.1).

(b) shows the relative levels of Luc activity, normalised to pPardC-luc, for the five transient transformation experiments (see Table 6.1).
3μg of genomic DNA from **LU352, FH3T1, FH3T3 and ES1T1** microplasmodia was digested with *HindIII* (section 2.3.1.3) and run on a 0.8% agarose gel with λ *HindIII* DNA size markers (section 2.3.2). The DNA in the gel was denatured, transferred to Hybond-N membrane and fixed to the membrane by exposure to UV (Sambrook *et al.*, 1989; section 2.4.1). The *PardC-hph* probe was radio-labelled (Feinberg and Vogelstein, 1983; section 2.4.4) and hybridised at high stringency (65°C) to the Southern blot. The blot was exposed to x-ray film at -80°C with an intensifying screen for 11 days. The relative positions of the λ *HindIII* DNA size markers is shown in bp. The image is a composite of different lanes from the same blot; irrelevant lanes were deleted.
CHAPTER 7: GENERAL DISCUSSION

Amoebae and plasmodia are very different cell types that differ in behaviour, cellular organisation and gene expression (sections 1.1, 1.2 and 1.3). Amoebae develop into plasmodia via an irreversible developmental transition that starts with an extended cell cycle (section 1.1.5). Time-lapse filming has shown that the extended cell cycle is an important time during both heterothallic and apogamic plasmodium development, since the cells become committed to plasmodium development, lose the ability to flagellate and begin to behave like plasmodia during this time (Bailey et al., 1987 and 1990). It is during the extended cell cycle that the changes in gene expression that result in the formation of a plasmodium are initiated (section 1.1.5). Northern blotting and immunofluorescence microscopy have shown that the amoeba-specific gene products are gradually lost over several cell cycles, while the plasmodium-specific gene products are gradually acquired (Sweeney et al., 1987; Solnica-Krezel et al., 1990 and 1991).

The isolation of amoebae which had lost the ability to develop into plasmodia following mutagenesis, led to the identification of several npf mutations which block apogamic development (Wheals, 1973; Anderson and Dee, 1977; Bailey et al., 1992a; Solnica-Krezel et al., 1995). Genetic analysis showed that most of these mutations were linked to matA, although some were not. The analysis of the npf mutations that were not linked to matA led to the idea that plasmodium development is controlled by several independent regulatory pathways that collectively result in the formation of a plasmodium (Bailey et al., 1992a; Burland et al., 1993b), and this notion was later confirmed by the analysis of double mutant strains (Solnica-Krezel et al., 1995; section 1.3.2). It is thought that matA initiates a cascade of gene action that results in the formation of a plasmodium (Bailey et al., 1995 and 1997).

In order to clone development specific genes that regulate plasmodium development, Bailey et al. (1992a) constructed a cDNA library from a culture which contained a high proportion of developing cells and enriched the library for genes that were expressed during development (section 1.3.3). The screening of this library led to the identification of a novel class of genes that were expressed primarily during development (section 1.3.3.3). These genes were called red (regulated in development) genes and the two that were discovered were called redA and redB (Bailey et al., 1999). Subsequent analyses showed that redA encodes a protein of unknown function, while redB encodes a protein that is related to invertebrate sarcoplasmic calcium binding proteins, although the function of these proteins is not clear (Bailey et al., 1999; section 1.3.3.3).
The main aim of this project was to identify and characterise more red genes since the analysis of such genes may lead to an understanding of how plasmodium development is controlled. I started by screening the subtracted cDNA library for more red genes. An examination of six cDNA clones led to the identification of two partial cDNA clones, D11/1100 and A18/1020, that appeared to represent red genes (section 3.2). These cDNAs were analysed further (Chapters 3 and 4).

Northern blotting and RT-PCR confirmed that D11/1100 represented a red gene; thus, the cDNA was renamed redE. Southern blotting showed that redE was a single-copy gene. The complete redE coding sequence, plus some of the upstream promoter sequence was obtained by cloning a genomic restriction fragment. Database searches showed that redE encoded a protein of unknown function, although the deduced protein sequence contained several potential phosphorylation and glycosylation sites. Thus, the function of RedE is unknown (Chapter 3).

Northern blotting confirmed that A18/1020 represented a red gene. Further northern analysis using RNA samples from heterothallic strains showed that the expression pattern of A18/1020 was similar in both apogamic and heterothallic development. This confirms the work of Bailey et al. (1999) which shows that the changes in gene expression that occur during the APT are similar in both types of development. Southern blotting suggested that A18/1020 was a single-copy gene that was a member of a gene family. Database searches showed that A18/1020 encoded part of the tail domain of a type II myosin heavy-chain protein. Thus, A18/1020 was renamed mynD (myosin developmental; Chapter 4).

Kohama et al. (1986) showed that P. polycephalum has two different type II myosin heavy-chain proteins (Section 1.2.2) and the work described in Chapter 4 suggests that even more exist (Bailey et al., 1999). The existence of multiple type II myosin heavy-chain proteins suggests that the isotypes perform different functions. The recent cloning of a full length cDNA encoding the major plasmodial myosin, mhcIIa, and a partial cDNA encoding the minor plasmodial myosin, mhcIIb, makes possible some interesting experiments (L. Nyitray, unpublished data: Chapter 4). If full length mhcIIb and mynD cDNA clones were obtained, mynD, mhcIIa and mhcIIb could be overexpressed in type II myosin heavy-chain deficient yeast cells (Rodriquez and Paterson, 1990). If the myosins did not correct the mutant phenotypes equally well, this would show that mhcIIa, mhcIIb and mynD are not functionally equivalent.

Immunolocalisation experiments could be used to localise the RedE and MynD proteins and pinpoint the expression of the genes to individual cells. The function of redE and
mynD could be investigated by gene knockout by homologous integration or by antisense experiments or by mis-expressing the genes (Chapter 6). Studies like these may lead to an understanding of the role that redE and mynD play in plasmodium development.

As discussed in Chapter 4, it seems likely that the red genes fall into several categories; some may be involved in regulating the changes in gene expression that occur during plasmodium development, while others, like mynD, may bring about the structural changes that occur during this time. Genes that fall into the former category could be identified using the strategy that is outlined below.

The putative red gene promoters could be cloned. The promoters would be defined by translationally fusing the putative promoter to the reporter gene, luc, stably transforming the construct into amoebae and assaying the transformant cells for Luc activity throughout development (Chapter 6). If the putative promoter contained all of the regulatory elements, the expression pattern of luc in the transformant cells would be the same as the expression pattern of the gene to which the promoter belongs in wild-type cells. Specific regulatory elements could then be identified within the promoters by either (i) looking for conserved domains or (ii) carrying out promoter deletions. The deleted promoters would then be translationally fused to the reporter gene, luc, and the construct would be transformed into amoebae, as above. If the deleted regions contained regulatory elements, the transformant cells would show an altered pattern of luc expression (Hori and Firtel, 1994). Gel retardation assays could then be carried out, using nuclear extracts from a culture of developing cells, to check that transcription factors bind to the regulatory regions (Carey, 1991).

The genes that encode the transcription factors could then be identified using the yeast one hybrid system (Vidal and Legrain, 1999). For example, the enhancer elements from the redE promoter could be linked to an inactive yeast promoter that is translationally fused to a selectable marker, and the construct would be stably transformed into yeast cells. Meanwhile, a cDNA library would be made in a yeast expression vector using RNA that was isolated from a P. polycephalum culture which contained a large proportion of developing cells. The cDNAs would then be transformed into the same yeast cells and only cells that contained the cDNA encoding the correct transcription factor would grow in the selective media. The cDNAs would then be isolated from the cells and sequenced. The function of any encoded transcription factors could then be examined by gene knockout or by mis-expressing the gene. Such studies may eventually lead to an understanding of how the regulatory cascade that is initiated by matA, results in the formation of a plasmodium.
Plasmodia are ideal for DNA replication studies because they contain millions of nuclei with naturally synchronous mitotic cycles. Density shift analysis and gene dosage experiments showed that the replication of the *red* genes was not confined to any compartment of the plasmodial S-phase; *redB* and *redE* are replicated during the first 30 min of S-phase, while *redA* is replicated between 30-60 min and 120-180 min from the onset of S-phase (Chapter 5). Studies have shown that the upstream promoters of genes that are abundantly transcribed in the plasmodium generally contain replication origins that fire at the onset of S-phase (section 1.6.1). The replication of *redB* and *redE* was examined to see if this is true of early replicating genes that are expressed at low levels in the plasmodium. 2D gel electrophoresis showed that both *redB* and *redE* are replicated at the onset of S-phase by forks arising from origins that lie either side of each gene (Chapter 5).

Since none of the *red* gene promoters contained origins that fire at the onset of S-phase, the data suggest a difference in the organisation of active origins in relation to genes that are expressed at high and low levels in the plasmodium. This work could be extended by examining the replication of the cell-type specific profilin genes in amoebae. The amoeba-specific profilin, *proA*, is not expressed in plasmodia, but is expressed in amoebae, while the plasmodium-specific profilin gene, *proP*, is not expressed in amoebae, but is expressed in plasmodia (Binette et al., 1990; Bailey et al., 1999). Pallotta et al. (1986) showed that the *proA* and *proP* mRNAs are very abundant; *proA* represented approximately 1.2% of the clones in an amoebal cDNA library, while *proP* represented approximately 1% of the cDNA clones in a plasmodial cDNA library. Bénard and Pierron (1992) examined the replication of *proA* and *proP* in the plasmodium and found that *proP* is tightly linked to a replication origin that fires at the onset of S-phase, while *proA* is replicated in mid S-phase. If the reverse is true in amoebae and *proA* is tightly linked to a replication origin that fires at the onset of S-phase, while *proP* is replicated at some later point, this would directly show that the replication timing of a gene depends on its transcriptional activity.
REFERENCES


